Characterisation and functional analysis of the putative *agr* system in *Clostridium acetobutylicum*

JAMIE SCOTT, BSc. (Hons), MSc.

Thesis submitted to the University of Nottingham
For the degree of Doctor of Philosophy

December 2012

Declaration

Unless otherwise acknowledged, the work presented in this thesis is my own. No part has been submitted for another degree in the University of Nottingham or any other institute of learning.

Jamie Scott

December 2012

Abstract

Clostridium acetobutylicum is an industrially important Gram positive organism which is capable of producing economically important chemicals in the Acetone, Butanol and Ethanol (ABE) process. An orthologue of the accessory gene regulator (*agr*) locus of *Staphylococcus aureus* has been found to be present in the genome of *Clostridium acetobutylicum*. In *S.aureus*, *agr* encodes a quorum sensing (QS) system that controls the expression of virulence in this species. Analysis of the *agr* region in *C.acetobutylicum* was conducted using reverse transcriptase PCR which showed the *agrB* and *agrD* genes to be linked. This was also the case with the *agrC* and *agrA* genes but there was no conclusive evidence to suggest that all 4 genes resided on the same operon. The use of *cat*-based reporter vectors which incorporate chloramphenicol acetyl transferase were used to look at the expression profiles of the *agrB* and *agrC* putative operons. The *agrC* construct showed activity consistent with the expected pattern of expression but this was not repeated with the *agrB* construct. Antisense RNA vectors were constructed with the intention of disrupting the *agr* genes but had no observable effect. This work was superseded by a newly available method to knockout the *agrA* gene by allelic exchange and the use of the ClosTron system to obtain gene inactivation in *agrB*, *agrC* and *agrA*. Gas chromatography analysis of these mutants showed little or no difference in product formation and a sporulation assay was developed which revealed that these mutants were inhibited in spore production. Finally, microarray analysis has been used to look at the effect of *agrB* inactivation on the gene expression of *C.acetobutylicum*. The expression of known sporulation genes was found to be differentially regulated. This study presents some of the first evidence to support the hypothesis that *agr* may be a major regulator in *C.acetobutylicum* and may act in a cell density dependent manner via a diffusible signal molecule.

Acknowledgements

I would like to thank my supervisors Professor Nigel Minton and Dr Alan Cockayne for their help and support throughout my PhD.

Many thanks to the *Clostridial* research group both past and present who have helped me along the way. Special thanks to Dr Glen Carter who taught me all his tricks in Molecular Biology and Dr John Heap who developed many of the techniques used in my work. I would also like to thank Victoria Wright for her time and hard work in producing the microarray data in this study and Professor Klaus Winzer for his help and advice.

I would like to thank the research group of Professor Philippe Soucaille at what is now National Institute for Applied Sciences, Toulouse, France who provided some of the plasmids and techniques used in my work. My thanks also to the research group of Professor Hubert Bahl in Rostock Germany who also provided advice on handling *Clostridium acetobutylicum*.

Many thanks to the staff of NHS Great Yarmouth and Waveney who have provided help and sympathy during the time I have spent writing my thesis.

Finally I would like to thank my family and friends who have provided so much support to me over the years. Thanks to my Grandparents in particular my Grandfather Derrick Lewis who was such a good role model and my Grandmother Mary Lewis whose faith in me was unwavering. Special thanks to my parents Glen Carter and Steve Carter for their hard work and support.

Last but not least I would like to thank my wife Lisa Scott for her patience support and love through the time of writing my thesis. Without her help and the support of those mentioned I would not have been able to complete this PhD.

Table of Contents

Table of Figures

Chapter 1

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of tables

Chapter 2

Chapter 5

List of Abbreviations

ABE	Acetone-butanol-ethanol fermentation
agr	Accessory gene regulator
AHL	*N*-Acyl homoserine lactone
AI	Autoinducer
AIP	Autoinducing peptide
Amp	Ampicillin
ARD	Alleviation of Restriction of DNA
AROS	Array-Ready Oligo Sets
ATP	Adenosine Triphosphate
bp	Base pair

BHFDR	Benjamini and Hochberg False Discovery rate
BHI	Brain Heart Infusion
CAT	Chloramphenicol Acetyltransferase
cDNA	Complementary Deoxyribonucleic acid
Cm	Chloramphenicol
CoA	Co-enzyme A
COG	Cluster of Orthologous Groups
CSF	Competence and sporulation factor
dCTP	Deoxycytidine triphosphate
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DPD	4,5- Dihydroxy 2,3- Pentanedione
DSF	Diffusible signal factor
dNTP	Deoxyribonucleoside 5’-triphosphate
DTNB	5,5'-Dithio-bis(2-Nitrobenzoic Acid)
EDTA	Ethylenediaminetetraacetic acid
Erm or ErmB	Erythromycin
FLP	Flippase Recombination Enzyme
FRT	Flippase Recognition Target
GTP	Guanosine-5'-triphosphate
gDNA	Genomic Deoxyribonucleic acid
h	Hour
HPK	Histidine Protein Kinase

HPLC	High-performance liquid chromatography
IEP	Intron Encoded Protein
IPTG	Isopropyl β-D-1-Thiogalactopyranoside
Kb	Kilobase
Kan or Km	Kanamycin
kV	Kilovolt
l	Litre
Lam	Lactobacillus *agr*-like module
MES	2-(N-morpholino)ethane sulphonic acid
mg	Milligram
min	Minute
ml	Millilitre
MLS	Macrolide-Lincosamide-Streptogramin antibiotic
mm	Millimetre
mM	Millimolar
mRNA	Messenger RNA
NADH	Nicotinamide Adenine Dinucleotide
ng	nanogram
ODn	Optical Density at wavelength n
ORF	Open Reading Frame
oriT	Transfer origin
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
PTS	Phosphotransferase System

QS	Quorum sensing
RACE	Rapid Amplification of cDNA Ends
RAM	Retrotranscription – Activated Selectable Marker
RAPs	Receptor Aspartyl Phosphatases
RBS	Ribosome Binding Site
RNA	Ribonucleic Acid
RNase	Ribonuclease
rpm	Revolutions per minute
RR	Response Regulator
RT	Reverse Transcriptase
RT-PCR	Reverse Transcription PCR
SAP	Shrimp Alkaline Phosphatase
sRNA	Small regulatory RNA
Ta	Annealing temperature
Tet	Tetracycline
Thm	Thiamphenicol
Tm	Melting temperature
TNB	5-Thio-2-Nitrobenzoic Acid
U	Arbitrary units
µF	Microfarads
UV	Ultraviolet
V	Volts
v/v	volume to volume
w/v	weight to volume
µg	microgram
µl	microlitre
µM	micromolar

X-gal	5-bromo-4-chloro-indolyl-β-D-galactopyranoside

1.0 INTRODUCTION

1.1 The Genus *Clostridium*

The genus *Clostridium* encompasses a group of Gram positive rod shaped bacteria. They belong to the division of low GC firmicutes and are obligate anaerobes. They are capable of producing endospores which enable members of the genus to survive extreme environments. The genus includes many significant pathogens including *Clostridium difficile*, *Clostridium botulinum*, *Clostridium perfringens* and *Clostridium tetani*. There are also a number of industrially relevant organisms within the genus. These include *Clostridium thermocellum* which is capable of digesting lignocellulose and producing ethanol.

1.2 *Clostridium acetobutylicum*

Clostridium acetobutylicum is capable of producing the industrially important solvents acetone and butanol as well as ethanol. It was first used in the ABE (acetone-butanol-ethanol) process to produce acetone for the production of cordite in the First World War. Chaim Weizmann is credited with the discovery of how the organism could produce these products using starch as a carbon source. For this reason *C.acetobutylicum* is sometimes called the “Weizmann organism”. The ABE process was used to produce these products until the cheap availability of oil made this process economically redundant. With the rising prices of oil, modern metabolic engineering and the increasing ability to genetically manipulate the organism the ABE process is becoming a more attractive way of obtaining these chemicals (Jones and Woods, 1986; Ezeji *et al.,* 2007).

The fermentation process of *C.acetobutylicum* is a biphasic process where the initial stage is acidogenic. Acid forming pathways are activated at this time which normally occurs during exponential growth (Andersch *et al.*, 1983; Hartmanis and Gatenbeck, 1984). The acids are reassimilated and

used in the production of the three solvents. The pathway involved in the formation of these products is shown in figure 1.1

1.3 Acidogenesis

The products acetate and butyrate are the two main acids produced at this stage of growth. The pathways involved in their production also produce ATP and so they are important in energy metabolism. More butyrate is produced than acetate at this stage as the formation of butyrate creates more of a balance to the redox equilibrium. This is because more NADH is consumed with this pathway. The concentrations of unassociated acids can reach a threshold where they become toxic to the cell and it is at this stage when solventogenesis is triggered. This transitional stage leads to a rapid change of gene expression (Durre, 2007).

During batch fermentation it has been found that if the concentration of undissociated acids reaches a threshold point of 57-60mmol.l^{-1} an “acid crash” occurs resulting in reduced levels of solvents (Maddox *et al.,* 2000) The chance of this occurring can be reduced by introducing appropriate buffering capacity to the medium or reducing the temperature of the fermentation (Maddox *et al.,* 2000; Bryant and Blascheck, 1988).

1.4 Solventogenesis

Commensurate with the fact that it is during the solventogenesis that the valuable products acetone and butanol are produced, the solventogenic phase of growth has been most studied. The typical ABE fermentation produces acetone, butanol and ethanol in the ratio 3:6:1. The solvents are useful in a wide variety of industries so a cheap and sustainable way of producing these products is being actively pursued.

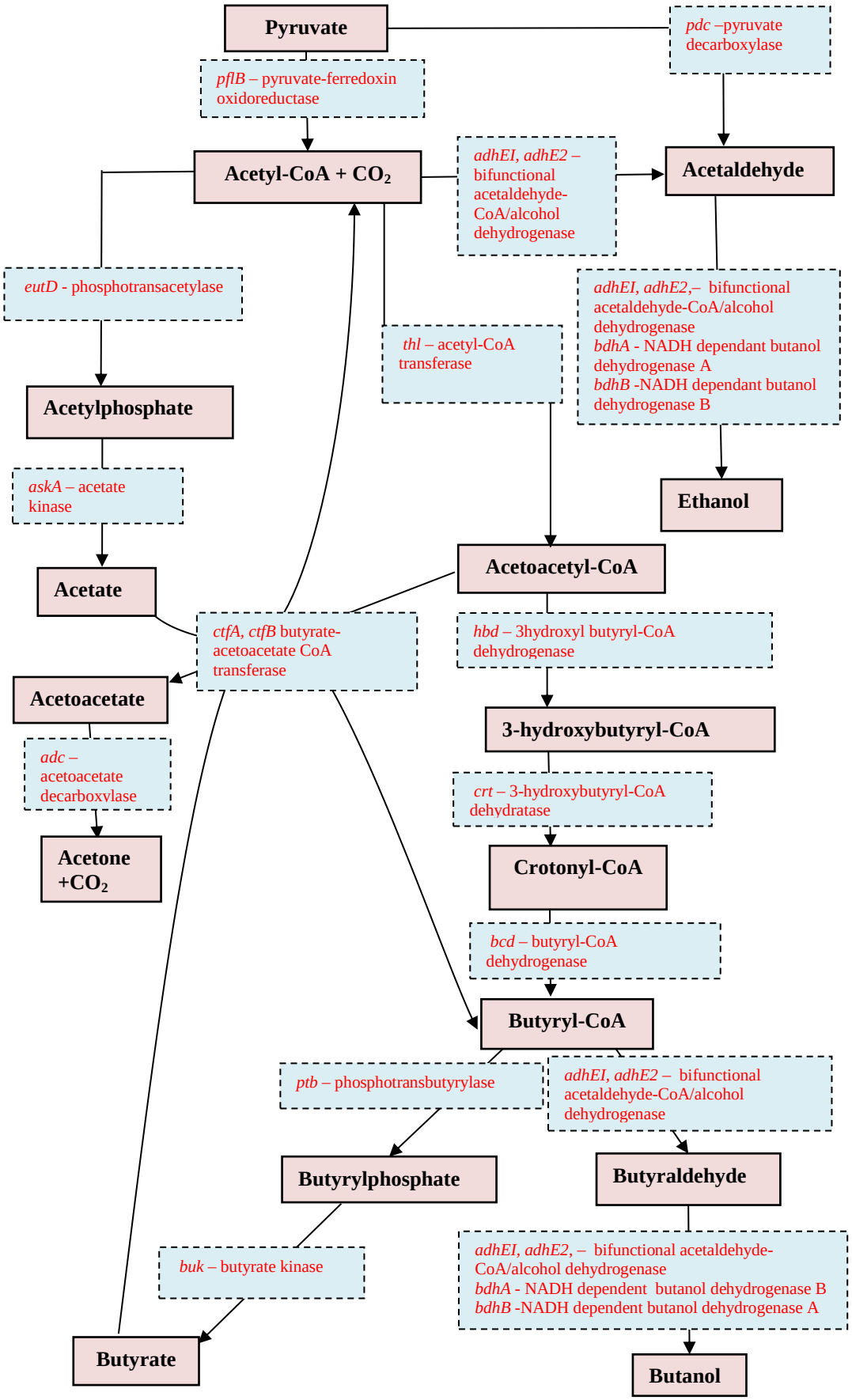


Figure 1.1 The biosynthetic pathways involved in acidogenesis and solventogenesis in *C.acetobutylicum*. Blue boxes show the relevant biosynthetic enzymes with their locus tags and pink boxes show the substrates and products.

1.4.1 Acetone

Acetone is a colourless flammable ketone. It was used in the 1st World War in the production of cordite for the munitions industry and the level of production went up during this time. Acetone has also been used in a variety of different areas including the plastic industry, as a cleaning agent, in paints and varnishes and as a storage agent for acetylene (Acetone – PubChem Public Chemical Database).

1.4.2 Butanol

This primary alcohol is a colourless liquid with a distinct odour. Half of butanol production is used in textiles, enamels and lacquers. It is also used as a solvent in the pharmaceutical industry and a component of brake fluids ("1-Butanol - Compound Summary". The PubChem Project. USA:National Center of Biotechnology Information; http://www.dow.com/productsafety/finder/nbut.htm)

There is renewed interest in butanol as a possible biofuel. It is superior to ethanol as a fuel additive as it has a higher energy content, lower volatility, is less corrosive and has less hydroscopicity meaning it is less likely to absorb water (Lee *et al.,* 2009). Furthermore, butanol is sufficiently similar enough to petrol to be used in the internal combustion engine.

1.5 Genes and Operons involved with Solventogenesis

The genes involved directly in the formation of solvents have been characterised. Many of these genes can be found on a large (size) megaplasmid called pSOL-1 (Cornillot *et al*., 1997). The key genes and their encoded enzymes are described in figure 1.1. A cluster of important genes reside in the *sol* operon shown in figure 1.2. Loss of pSOL-1 results in the loss of the ability to produce solvents.

The genes involved in the formation of acetate are formed into an operon in the order *pta-ack*. The genes involved in the formation of butyryl-CoA are clustered in the *bcs* operon in the order *crt-bcd-etfB-etfA-hbd*. *etfA* and

etfB encode gene products with homology to electron transfer flavoproteins (Bennett and Rudolph, 1995; Boynton *et al.,* 1996). The genes involved in butyrate synthesis, *ptb* and *buk* also form an operon (Cary *et al.,* 1988).

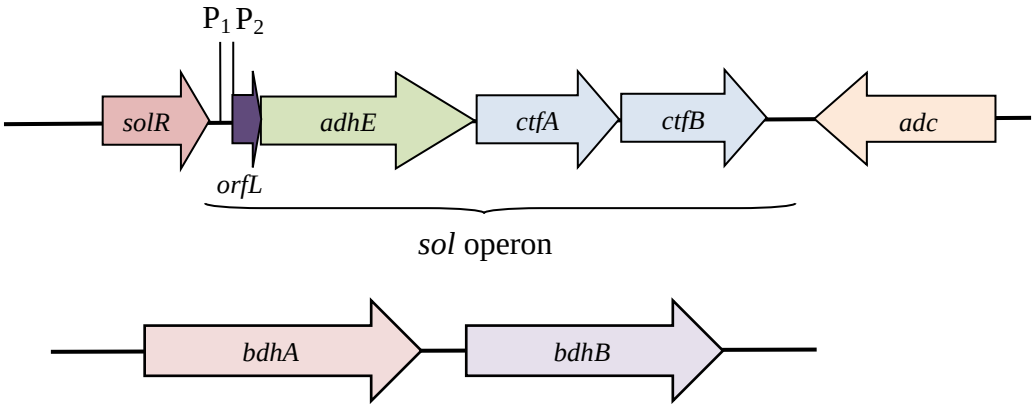


Figure 1.2. The *sol* operon in *C.acetobutylicum* and the *adc* gene which reside on the pSOL-1 megaplasmid and the two bdh genes which reside on the chromosome which are involved in solventogenisis**.**

Of the genes described in figure 1.1, the *adc* gene encoding acetoacetate decarboxylase is present in *C.acetobutylicum* as a monocistronic gene (Gerischer and Durre 1990; Petersen *et al.,* 1993).

The *sol* operon consists of *ctfA* and *ctfB*, (the 2 subunits of butyrate/acetoacetate-CoA transferase) and *adhE,* one of the 2 acetaldehyde-CoA dehydrogenases found in *C.acetobutylicum*. Also present upstream of *adhE* is *orfL* which encodes a small peptide of as yet unknown function. The *sol* operon was found to have 2 promoters, P_1 and P_2. Reporter studies have shown that P_1 is the true promoter and that P_2 is an RNA processing site (Thormann *et al.,* 2002).

A second *adhE* gene in *C.acetobutylicum* (CAC0035) has been found. As such, this *Clostridium* is the first *Clostridium* species to have been found to contain two *adhE* genes (Fontaine *et al.,* 2002). This second gene, *adhE2* was found to be expressed only in “alcohologenic” continuous culture

where no acetone is produced as a result of specific substrate combinations of glycerol and glucose (Fontaine *et al.,* 2002).

The *adhE2* gene is found on pSOL-1 along with the *adc-sol* locus. Two other genes encoding butanol dehydrogenases, *bdhA* and *bdhB* are arranged contiguously on the chromosome (see figure 1.2).

1.5.1 Expression profiles of metabolic genes

The expression profiles of the metabolic genes of *C.acetobutylicum* have been studied by DNA microarray analysis (Alsaker and Papoutsakis, 2005). The expression of the genes involved in acetate and butyrate biosynthesis was found to peak at the transitional stage of growth. Acetate genes had the highest fold up-regulation. The expression of the *bcs* (butyryl-CoA synthesis) operon was not surprisingly found to increase 1.9 fold in the stationary phase of growth. The *thl* gene (encoding Thiolase) followed a similar expression pattern. The *bdh* genes were found to be up-regulated at a time point coinciding with the onset of sporulation. There was found to be differential expression between the 2 genes with *bdhA* having much lower levels of expression compared with that of *bdhB*.

The genes belonging to the *sol* operon and *adc,* as expected, were found to be up-regulated in a similar pattern to the *bdh* genes.

1.6 Regulation of Solventogenesis

Much is known about the genes involved in acidogenesis and solventogenesis in *C.acetobutylicum*. Until recently there has been very little knowledge of the regulatory mechanisms involved in their expression and of the transition between the two stages of growth.

1.6.1 Spo0A

This regulatory factor was originally identified in *Bacillus subtilis* and is a key regulator of the sporulation pathway in this organism (Ferrari *et al.,* 1985). In *B.subtilis* Spo0A is phosphorylated by Spo0B which is the last component of a characterised phosphorylation cascade. The activated Spo0A can then bind to a consensus sequence known as a 0A box (Hoch, 1993a; Hoch, 1993b). A recent study identified the most probable consensus sequence as 5′-TTTGTCGAAAA-3′ using two different methods of analysis and found 121 genes directly up-regulated by Spo0A in *B.subtilis*.

A homologue of Spo0A has been found in *C.acetobutylicum* and was found to affect the expression of genes at the transitional phase between acidogenic to solventogenic phases of growth. The strain SKO1, which has its *spo0A* gene disrupted, had lower expression of the *sol* operon and *adc* (Harris *et al.,* 2002). The phosphorelay system which activates Spo0A in *B.subtilis* is missing in clostridia and there have been various studies into what might activate the clostridial Spo0A (Alsaker and Papoutsakis, 2005; Jones *et al.,* 2008). Studies into overexpression of *spo0A* revealed 123 genes which had their expression altered (Alsacker *et al.,* 2004).

1.6.2 Abr and SinR

Other transcription factors have been shown to act in concert with Spo0A in *B.subtilis*, including the AbrB and SinR proteins (Grossman, 1995). AbrB acts as an inhibitor of sporulation in effect acting as an opposite regulator to Spo0A. It modulates catabolism during carbon limitation and acts as a positive and negative regulator of competence during cell growth (Fisher *et al.,* 1994; Hahn *et al.,* 1995). *C.acetobutylicum* carries 3 homologues of Abr (CAC0310, CAC1941 and CAC3641). The use of reporter genes has, however, established that CAC0310 is the true AbrB functional homologue (Scotcher *et al.,* 2005). Expression of CAC0310 is highest during the transitional phase of growth and disruption of CAC0310

delayed solventogenesis and sporulation. There is also a homologue of *sinR* in *C.acetobutylicum* (Scotcher *et al.,* 2005). This helix-turn-helix DNA binding protein is synthesised during vegetative growth and sporulation in *B.subtilis*. It can form dimers *in vivo* which directly inhibit sporulation initiators including Spo0A (Cervin *et al.,* 1998; Mandic-Mulec *et al.,* 1995). SinR itself is regulated by SinI which inhibits *sinR* activity at the onset of sporulation (Bai *et al.,* 1993). *sinR* and *sinI* form a bicistronic operon (shown in figure 1.3).

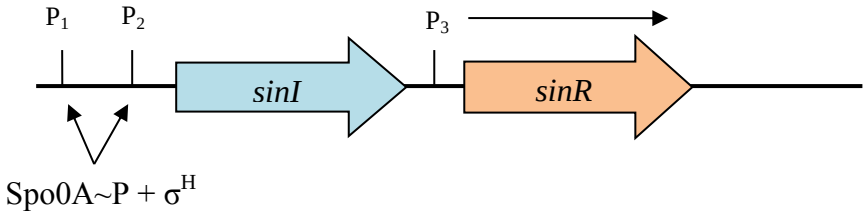


Figure 1.3. Illustration of the *sinI/R* operon of *B.subtilis* showing all three promoters. P3 is responsible for the continuous expression of SinR. The P1 and P2 promoters require activated Spo0A and SigH for expression (Gaur *et al.,*1986; Gaur *et al.,* 1988).

CAC0541, the homologue of *sinR* in *C.acetobutylicum,* was found to be expressed weakly throughout a 24 h growth curve. It was disrupted through antisense RNA but this was found not to effect solventogenesis (Scotcher *et al.,* 2005).

1.6.3 SolR

SolR is a possible transcriptional repressor of solventogenesis with a helix-turn-helix DNA binding motif (Nair *et al.,* 1999). Inactivation of *solR* was found to result in higher levels of butanol and acetone while overexpression of *solR* resulted in the opposite effect (Nair *et al.,* 1999; Harris *et al.,* 2001). Further studies found that SolR did not bind to any DNA upstream of *sol* and overexpression of *solR* in *E.coli* revealed a different pattern of glycosylated exoproteins to wild type (Thormann and Durre, 2001). This suggests that SolR is not a direct repressor of solventogenesis genes. Furthermore an explanation has been given for

the effect of the overexpression of *solR* on the reduction of solvents produced. It was suggested that the plasmid expressing *solR* had a Spo0A binding site and so the titration of Spo0A itself could have caused this effect (Thormann and Durre, 2001). The fact that a knockout of *solR* did increase solvent production does suggest it has some yet to be elucidated role in solvent formation.

1.7 Metabolic engineering of *C.acetobutylicum*

Over the years there has been extensive interest in developing methods to improve the yields of solvents from *C.acetobutylicum* and to make production, particularly of butanol, economically viable. Three main routes of metabolic engineering are being pursued in order to achieve this goal. The first is to increase solvent production directly by altering metabolic pathways and blocking sporulation, the second is to increase solvent tolerance and the third method is to enable *C.acetobutylicum* to utilise cheaper substrates.

One of the limitations of metabolic engineering with clostridia is the lack of genetic tools available. The presence of the Cac824I restriction system in *C.acetobutylicum* initially prevented the transfer of recombinant vectors prepared in *E.coli*. This was at first overcome by using shuttle vectors such as pFNK-1 which lacks the recognition sites of Cac824I (Mermelstein *et al.,* 1992) and thereafter with the use of an *E.coli* donor carrying the vector pAN1 (Mermelstein and Papoutsakis, 1993), and more recently pAN2 (Heap *et al.,* 2007), which contain the φ3TI methyl transferase gene of the *B.subtilis* phage φ3tI. This protects the plasmid from the restriction by Cac824I.

Studies have been conducted using plasmids to overexpress genes involved with solventogenesis. The overexpression of *spo0A* has resulted in increased levels of butanol compared with a plasmid only control strain (Harris *et al.,* 2002). Overexpression of *adc*, *ctfA* and *ctfB* genes has

resulted in 90% and 37% higher levels of acetone and butanol, respectively, than the plasmid control strain (Mermelstein *et al.,* 1993).

Progress has also been hampered by the inability to produce knockouts not just in *C.acetobutylicum* but in the whole genus *Clostridium*. Until recently, only a handful of gene knockouts had been made. Four of these were derived through the unstable insertion of a plasmid by a single crossover event and were obtained by labour intensive screening of thousands of transformants (Shimizu *et al.,* 1994; Wilkinson and Young, 1994; Green and Bennett, 1996; Green *et al.,* 1996; Sarker *et al.,* 1999; Liyanage *et al.,* 2001; Harris *et al.,* 2002; Huang *et al.,* 2004; O'Connor *et al.,* 2006; Raju *et al.,* 2006).

1.7.1 Reporter Studies

As a result of the difficulty of achieving gene knockouts in *C.acetobutylicum* much of the investigation into the physiology of the organism has employed the use of other genetic tools. These include the use of reporter based systems and antisense RNA technology. The use of reporter systems has made it possible to study the expression of genes in *C.acetobutylicum*. A total of 4 genes have been used as reporters in this organism including *lacZ*, *lucB*, *gusA* and *catP*. The *lacZ* reporter system was developed using a gene from *Thermoanaerobacterium thermosulfurogenes* and used to analyse the expression of *ptb*, *thl*, *adc*, *bdhA*, *bdhB*, *sol*, and one of the *abr* homologues in *C.acetobutylicum* (Tummala *et al.,* 1999; Feustel *et al.,* 2004; Scotcher and Bennet, 2008). It has also been used in the previously mentioned *solR* study (Thormann *et al.,* 2002). The Feustel study (Feustel *et al.,* 2004) also used a bioluminescence gene *lucB* from *Photinus pyralis* encoding luciferase as an alternative reporter to *lacZ* to analyse the expression of the *sol* and *ptb-buk* operons. The *gusA* gene from *E.coli* has been used as a reporter system in *C.acetobutylicum* to study expression of the *thlA* promoter (Girbal *et al.,* 2003). The *catP* encodes a chloramphenicol acetyl transferase which can confer resistance to this antibiotic. This reporter

system has been used to investigate the promoter activity of phospholipase C and α toxin in *C.perfringens* (Matsushita *et al.,* 1994; Bullifent *et al.,* 1995). These reporter systems have been an important development in facilitating metabolic engineering in *C.acetobutylicum*.

1.7.2 Antisense RNA

Antisense RNA technology has been used successfully to alter solvent levels in *C.acetobutylicum*. This tool relies on the introduction of a complementary strand of RNA to the target gene which inhibits its translation. This mechanism of down-regulation has been used to try and increase the ratio of butanol to acetone in *C.acetobutylicum*. A study by Desai *et al,* (1999) produced a strain with antisense RNA directed against the *buk* gene. This resulted in a reduction in the activity of Buk by 80-90% compared to the control. This strain was found to produce 50% and 30% higher levels of acetone and butanol, respectively. The levels of Ptb activity were also impaired, which is probably due to the fact they share the same mRNA transcript which was subject to decay. The study also produced a strain with antisense RNA directed against *ptb*. This strain showed 50-80% lower Ptb specific activity and also 30-60% lower Buk activity, again attributed to the fact that the two genes share the same RNA transcript. Solvent formation was found to be lower than in the control strain with increased levels of lactate produced relative to the wild type strain. No effect on butyrate formation was identified in either strain which could possibly be explained by the fact that butyrate synthesis enzyme levels do not control butyrate production. This metabolic pathway is necessary for energy production and is, therefore probably strictly controlled. A second metabolic engineering study involved the creation of a strain producing antisense RNA directed towards the *ctfB* gene of the *sol* operon (Tummala *et al.,* 2003a). This resulted in lower levels of solvent production indicating that a reduction in *ctfB* had an effect on the operon as a whole. As a result the strategy of down regulation of *ctfB* along with the overexpression of *adhE* was employed in order to get a desirable ratio of acetone to butanol production. The ratio increased with this strain compared to the control but final butanol levels were lower than the wild

type strain. This was explained by the fact that production of butanol needs a functioning alcohol/aldehyde dehydrogenase encoded by the *adhE* gene to convert butyrate to butyryl-CoA which feeds butanol and ethanol production. The strain was also not unexpectedly found to produce very high levels of ethanol. Because the *ctfB* and *adhE* genes reside on the same mRNA transcript, the *ctfB* asRNA resulted in a down regulation of both the *ctfB* and *adhE* genes thus resulting in lower butanol production. The same group looked at using antisense to block the *adc* gene in an effort to reduce the amount of acetone (Tummala *et al.,* 2003b). This was found not to reduce acetone levels. Three antisense RNA strains targeted to *ctfA* and *ctfB* encoding subunits and both *ctfA* and *ctfB* together were found to reduce acetone levels suggesting that this enzyme was the rate limiting step in the acetone formation pathway. A later study combined *adhE* and thiolase overexpression to increase alcohol yields (Sillers *et al.,* 2009). However, the production of antisense RNA directed against *ctfB* in combination with thiolase overexpression did not lead to an altered product formation indicating that even more complex methods of metabolic engineering will be necessary to achieve high butanol yields.

These studies illustrate the difficulties encountered in metabolic engineering to increase solvent yield. Antisense techniques have so far revealed a limited amount of information on the metabolic pathways involved in solventogenesis. Achieving regular knockout of genes will develop knowledge in this area further. Overexpression of proteins can only have a limited effect on the organism as solvents, especially butanol are toxic to the cell. An efficient method of product removal or achieving a greater tolerance to solvents is another method most likely to be additionally necessary if increased productivity is to be obtained. One of the initial studies directed at increasing solvent tolerance was conducted where the *cfa* gene was overexpressed (Zhao *et al.,* 2003). This gene is involved in the formation of cyclopropane fatty acids at the expense of unsaturated chains in the membrane. This is thought to confer greater resistance to solvents. Overexpression of this gene resulted in lower levels of solvent formation compared to the control strain. The reason for this

could be because this gene also confers resistance to acids formed by the organism which enables less of the metabolic resources to be devoted to acid assimilation and subsequent solvent production. There have been several studies looking at genes involved with the stress response which may play a role in tolerance to solvents. The accumulation of solvents in the cellular environment has been shown to increase membrane permeability leading to the efflux of ATP, protons, ions and macromolecules (Sikkema *et al.,* 1995). In *C.acetobutylicum* specifically, high butanol concentration has been known to disrupt the cell by inhibiting nutrient transport, glucose uptake and membrane bound ATPase activity (Bowles and Ellefson, 1985). Heat shock proteins, or molecular chaperones, play an important role in normal cellular growth in protein synthesis, transport and folding of proteins. During shock and stress conditions they may primarily serve to prevent aggregation and assist in protein folding (Narberhaus *et al.,* 1992). The overexpression of one of these proteins, *groESL*, resulted in increased yields of solvents when compared to a control strain (Tomas *et al.,* 2003a). Other heat shock genes have been implicated in solvent tolerance including *dnaK* (Narberhaus *et al.,* 1992). The use of a genomic library with transcriptional array using butanol stress has also highlighted similar genes (Borden and Papoutsakis, 2007).

1.7.3 Microarray

The use of microarray techniques has greatly increased the knowledge of *C.acetobutylicum* gene expression. This powerful technique can be used to look at the regulation of a great number of genes over the normal range of growth or compare the expression of genes over time under different conditions. Combined with gene overexpression, antisense RNA technology or gene knockout, it can be a useful tool to determine the regulatory effect of specific genes. The majority of this microarray work has been done by the Papoutsakis group currently at the University of Delaware (Alsaker *et al*., 2004; Alsaker *et al*., 2005; Borden and Papoutsakis 2007; Jones *et al*., 2008; Parades *et al*., 2004; Tomas *et al*., 2003a; Tomas *et al*., 2003b; Tummala *et al*., 2003a; Tummala *et al*.,

2003c) . Some of these studies have already been mentioned previously, with many studies also focusing on the underlying genetic control of sporulation.

1.8 Sporulation

After solventogenesis, sporulation is perhaps the second most studied physiological event in *C.acetobutylicum*. It is linked to solventogenesis but is not exclusive to this event (Scotcher and Bennett, 2005). The stages of sporulation and its regulatory events have been most extensively studied in *B.subtilis.* The 8 stages of sporulation are illustrated in figure 1.4. The vegetative state of cell division is considered to be stage 0 in the sporulation pathway. This is followed by stage I where two copies of the chromosome condense and form a filament along the long axis of the cell. Asymmetric division then occurs when the cell divides at a sub polar site resulting in two cells of unequal size. The two cells are termed the mother cell and the prespore (often described as the forespore). Initially one third of the chromosome is present in the prespore. The rest of the chromosome is transported to the prespore by SpoIIE, a DNA translocase (Wu and Errington 1998; Bath *et al.,* 2000).

The next stage involves the engulfment of the prespore where there is septal thinning and migration of the septal membrane around both sides of the prespore. The membranes then fuse at the cell pole in effect releasing the prespore within the mother cell. The prespore is, therefore, covered by two membranes at this stage, derived from each of the two cells. Deposition of two peptidoglycan layers, the primordial germ cell wall and the cortex takes place between these two membranes in stage IV. This stage is followed by the formation of the spore coat which is a complex formation of proteins. Stage VI involves the maturation of the spore within the mother cell where it gains its resistant properties after which the mother cell lyses and releases the spore at stage VII. Gene's which have been found to be involved with sporulation are usually designated by a number indicating at which stage in sporulation the gene's mutation

causes a blockage, and a distinguishing letter designation (Piggot and Coote, 1976).

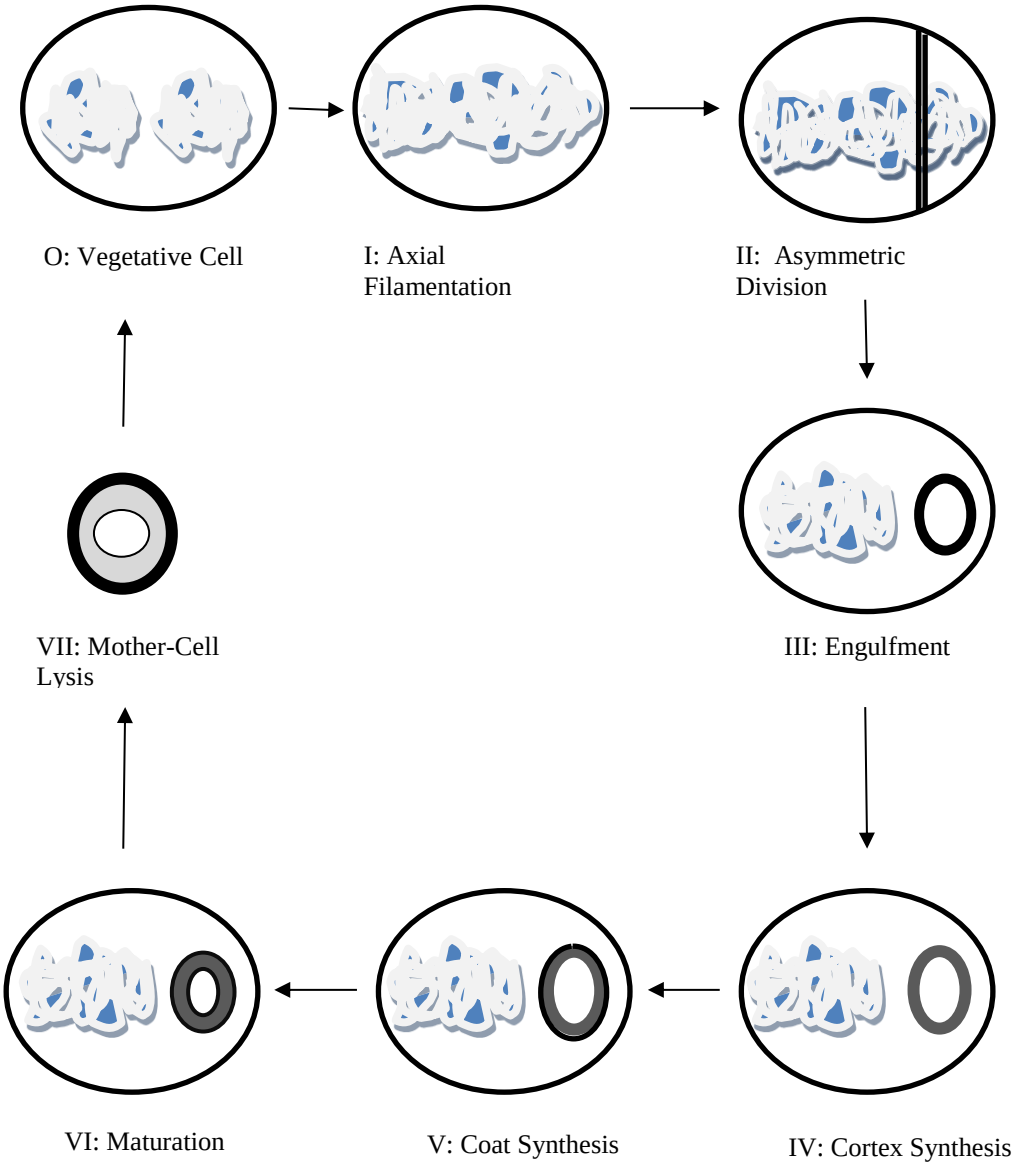


Figure 1.4 The *B.subtilis* sporulation pathway showing the 8 stages of sporulation**.** Stage 0 represents normal vegetatively dividing cells with two chromosomes (represented by disorderly lines within the cell). Stage I includes the start of axial filimentation and was not recognised universally as mutants were not initially isolated which arrested sporulation at this stage. A *citC* mutant has been isolated which blocks sporulation at this stage which provides evidence that this is a valid stage (Jin et al,. 1997) This then forms a continuous structure along the axis of the cell as represented in stage II. The single line is a basic representation of asymmetric division. Stage III denotes the completion of engulfment where the septal membrane undergoes autolysis and the points of attachment to the peripheral membrane migrate to the cell pole. The engulfment is complete when the membranes fuse resulting in a detached prespore surrounded by the mother cell. The synthesis of the germ cell wall and cortex then occurs in stage IV shown by the grey circle, followed by the spore coat represented by a dark outer circle at stage V. Dehydration of the prespore also occurs at this stage resulting in a 'phase bright' appearance. Stage VI denotes the maturation of the spore giving its full resistant properties. Thereafter the mother cell can lyse, releasing the spore at stage VII.

1.8.1 Genetic control of sporulation

The genetic basis of sporulation in *C.acetobutylicum* is little understood in comparison to *B.subtilis*. As a result much of the information given regarding this physiological event will be with regards to the *B.subtilis* paradigm. There are, however, many similarities between the *B.subtilis* model and sporulation in *C.acetobutylicum,* which shares many homologues of *B.subtilis spo* gene's and sporulation specific sigma factors.

1.8.2 Role of Spo0A

As already stated, the *spo0A* gene is involved with solventogenesis in *C.acetobutylicum* but as its name suggests it is also primarily involved in sporulation. It is considered the master sporulation regulator in *B.subtilis* (Hoch *et al.,* 1993a; Hoch *et al.,* 1993b). In this organism activation of Spo0A by phosphorylation is achieved through the participation of a phosphorelay system described in figure 1.5. The phosphorylated form of Spo0A (~P) then goes on to activate sporulation specific genes, including major sporulation dependent sigma factors. An equivalent to the phosphorylation cascade which activates Spo0A has not been found in *C.acetobutylicum* although some potential candidate proteins are being studied which may phosphorylate Spo0A (Jones *et al.,* 2008). This study explored the possibility of Spo0A being directly phosphorylated by an orphan histidine kinase. In *C.botulinum* an orphan kinase CBO1120 is able to phosphorylate Spo0A. Jones *et al.,* 2008 identified 6 orphan kinases including CheA, a response regulator which could be implicated in phosphorylating Spo0A. The expression profiles of each of these genes were analysed to see if their expression peaked just before or during the activation of Spo0A. Spo0A activity was measured using the expression of the *sol* operon as an indicator. Of the six kinases, CAC1437 and CAC0903 showed most similarity to the expression profile
of CBO1120. The study also looked at the possibility that a non-orphan two component response system could be implicated in activating Spo0A.

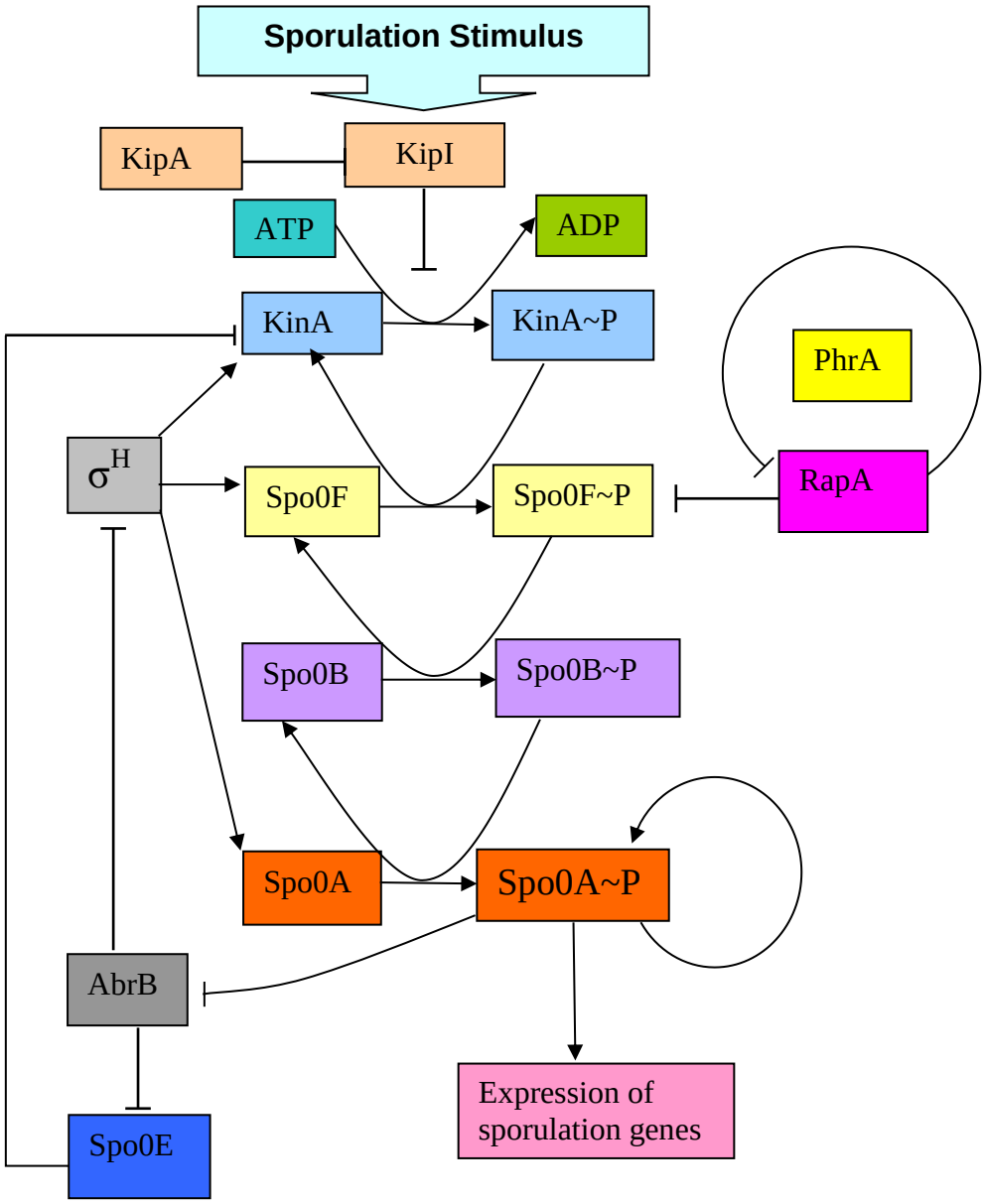


Figure 1.5 Phosphorylation cascade leading to the activation of Spo0A in *B.subtilis***.** Lines with arrows denote positive regulation and lines with bars at the end denote negative regulation. Spo0A is phosphorylated by Spo0B~P which in turn is phosphorylated by Spo0F~P. Spo0F is itself phosphorylated by at least 3 kinases KinA, KinB and KinC (Stephensen and Hoch, 2002). Here KinA is shown. KipI was identified as a protein which when overexpressed inhibited sporulation in a KinA dependent manner. The product of *kip*A, found adjacent to *kip*I was found to counteract this effect (Wang L *et al.,* 1997). Rap (represented here by RapA) proteins have been found to directly inhibit Spo0F phosphorylation which is in turn regulated by Phr proteins in a density dependent manner. This regulation will be mentioned in further detail later in the chapter.

The 30 remaining annotated histidine kinases were investigated to see if they showed upregulation just before the induction of the *sol* operon. Six

genes were found to peak in expression at 8 h of growth, just before *sol* induction. Of these six genes, four, CAC0225, CAC0863, CAC1582 and CAC2434, maintained variable levels of expression after this time point. In *B.subtilis spo0A* is itself activated by σ^H, a vegetative state sigma factor which also activates genes involved with Spo0A phosphorylation, such as *kin*A and *spo0F*.

Activated Spo0A, as well as promoting its own expression in *B.subtilis* represses the expression of *abrB,* a gene which produces a product involved in repressing stationary phase events (Robertson *et al.,* 1989). Knocking out an *abrB* homologue in *C.acetobutylicum* was found to increase acid production without effecting spore morphology (Scotcher and Bennett, 2005).

1.8.3 Axial filament formation and asymmetric division

The first stages of sporulation occur when the cell divides asymmetrically. This is preceded by axial filament formation. This stage of growth has still to be defined and no *spo0* mutations have been found which specifically prevent this stage. This may indicate that it is not a fixed event in sporulation (Piggot and Coote, 1976). However, it has been found that asymmetric division does not occur without axial filament formation, suggesting that this checkpoint functions in order to couple the two events (Graumann and Losick, 2001). Some of the proteins thought to be involved in axial filament formation include DivIVA and RacA. DivIVA is a coiled-coil, tropomyosin-like protein and a functional homologue of MinE. It sequesters the MinCD division inhibitor to the cell's poles which directs division to the cell's mid-point during vegetative growth (Cha and Stewart 1997; Edwards and Errington, 1997). The protein may be involved with positioning the *oriC* region of the chromosome to the cell pole during sporulation (Thomaides *et al.,* 2001). This study isolated mutants of *divIVA* and found some to be specifically defective in sporulation with compartments devoid of DNA. RacA is a protein which has been proposed to bind DivIVA and displace the division inhibitor MinCD (Ben-Yehuda *et al.,* 2003).

Once asymmetric division has taken place, the genetic regulation of sporulation is in effect compartmentalized with differences in gene expression in the prespore and mother cell. During asymmetric division Spo0A triggers the formation of a Z ring at both poles of the cell (Levin and Losick, 1996). The essential tubulin homologue FtsZ forms a ring under normal vegetative conditions where division occurs (Wang and Lutkenhaus, 1993). The *ftsAZ* operon is involved at the asymmetric phase of growth and its expression is thought to be modulated by σ^H and *spoIIE* (Gholamhoseiniam *et al.,* 1992; Gonzy-Treboul *et al.,* 1992; Ben-Yehuda and Losick 2002; Khvorova *et al.,* 1998). SpoIIE localises to the asymmetric division sites in a FtsZ-dependent manner (Levin *et al.,* 1997). The mechanism by which these two proteins interact is still unclear.

In order for sporulation to proceed further there needs to be a transfer of the remaining DNA into the prespore from the mother cell. The protein involved with this mechanism is SpoIIIE. Mutation of SpoIIIE causes the distal two thirds of the chromosome to be trapped in the mother cell, thereby halting sporulation (Wu and Errington, 1998).

1.8.4 Sigma factors involved in sporulation

The sporulation pathway in *B.subtilis* is controlled by a variety of sigma factors as shown in figure 1.6. Asymmetric division leads to the activation of the first prespore specific sigma factor σ^F. This is followed by activation of σ^E in the mother cell. The completion of engulfment leads to the activation of a second prespore specific sigma factor σ^G and the mother cell specific σ^K. The expression of each sigma factor is co-ordinated between the prespore and the mother cell by a series of biochemical signals.

1.8.5 Role of σ^F

The activation of the first prespore specific sigma factor is controlled by the *spoIIA* operon. σ^F is only active in the prespore. It is initially kept inactive by the binding of SpoIIAB and is released upon binding by SpoIIAA. SpoIIAA is itself phosphorylated by SpoIIAB (Min *et al.,* 1993) and is reactivated by SpoIIE (Duncan *et al,* 1995). σ^F is involved in the regulation of SpoIIR which is involved in the activation of σ^E (Karow *et al.,* 1995), SpoIIIG which is involved in the regulation of the late prespore specific σ^G (Sun *et al.,* 1989) and SpoIVB, a regulator of the late mother cell specific σ^K.

1.8.6 Role of σ^E

This sigma factor is transcribed in an inactive form in the mother cell. It is one of the two products of the *spoIIG* operon. *spoIIGB* encodes the pro-σ^E which is processed into an active form by SpoIIGA. SpoIIGB was found to be expressed in a Spo0A-dependent manner and also requires σ^A RNA polymerase (Wang *et al.,* 2007). The SpoIIGA is itself activated in response to SpoIIR which as mentioned is regulated by σ^F.

σ^E regulates important genes involved in the engulfment process. These include *spoIID*, *spoIIM* and *spoIIP* which are also involved in preventing a second asymmetric division from occurring in the mother cell (Abanes-de Mello *et al.,* 2002; Coote, 1972; Eichenberger *et al.,* 2001; Piggot and Coote, 1976; Pogliano *et al.,* 1999; Rong *et al.,* 1986; Smith *et al.,* 1993; Smith and Youngman, 1993). σ^E also directs expression of the *spoIIIA* operon which is required for prespore stability (Doan *et al.,* 2009). σ^E is also required for the expression of the late mother cell specific σ^K (Kunkel *et al.,* 1990; Stragier *et al.,* 1989).

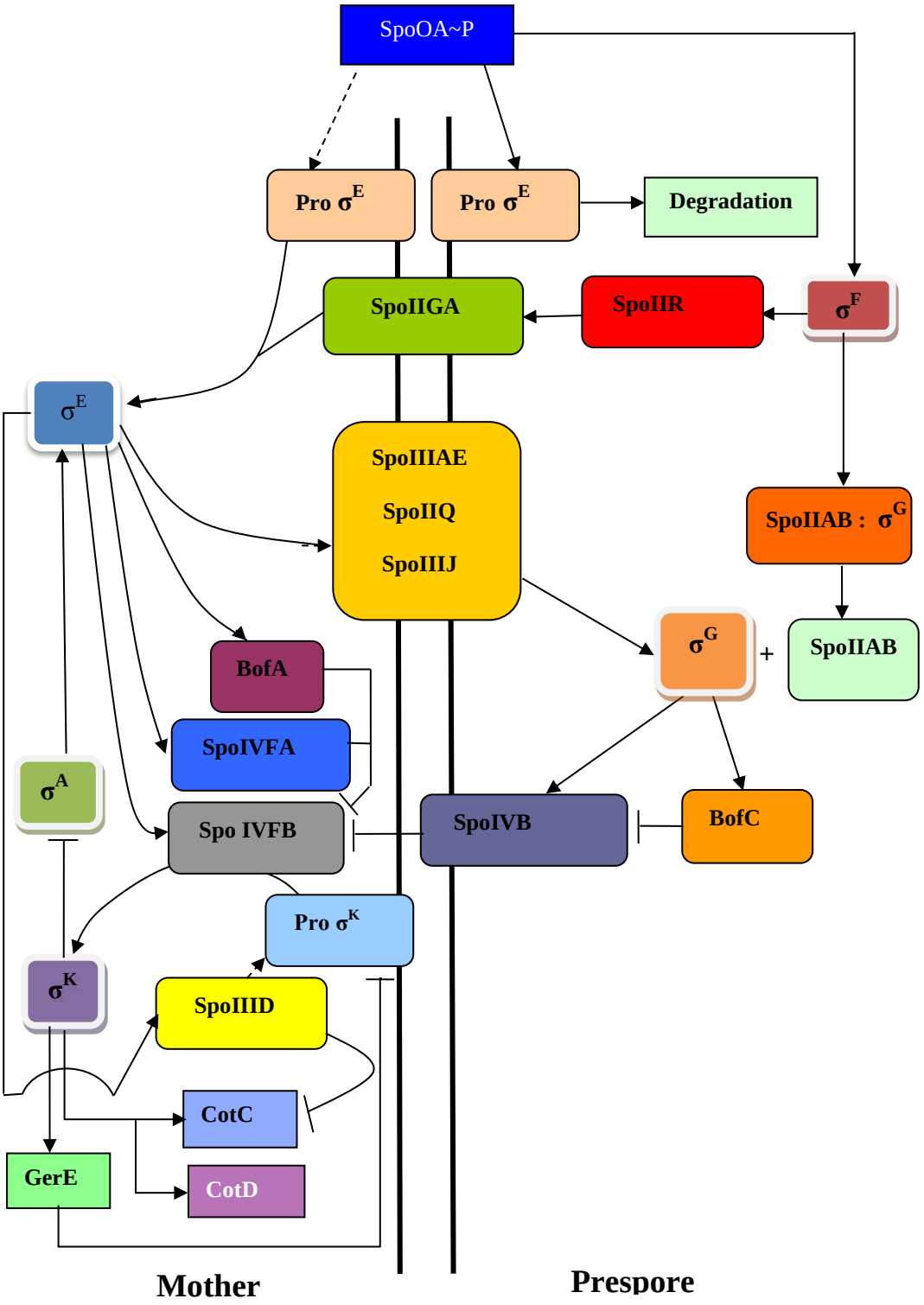


Figure 1.6. A simplified representation of the role of sigma factors and other regulatory proteins involved in the regulation of sporulation in *B.subtilis*. Arrows denote positive regulation and barred lines denote negative regulation

1.8.7 Engulfment (σ^G)

The process of engulfment is the second major stage in sporulation. The late stage prespore specific σ^G is activated after engulfment is completed. As mentioned previously, it is under the control of σ^F (Sun *et al.,* 1991). Once it has been synthesised it regulates its own expression in a positive feedback loop. Originally it was thought that σ^G activity was controlled by the *spoIIIA* operon directly with SpoIIQ (Londono-Vallejo *et al.,* 1997). Doan *et al.* (2009) demonstrated that SpoIIQ and the SpoIIIA seem to form a multimeric membrane complex which spans the double membrane surrounding the prespore. The study suggested that this structure maintains prespore integrity and possibly provides the mechanism by which the main cell can nurture the prespore. Therefore, these products may not actually be needed for σ^G activity. SpoIIIJ, which has also been implicated inσ^G activation, has also been found to interact with SpoIIIAE (Serrano *et al.,* 2008). Further work needs to be done in order to ascertain if σ^G actually requires processing in order to become active.

σ^G regulates many genes involved in sporulation, germination and protection of the spore. It also regulates the production of SpoIVB which signals the mother cell to activate σ^K. Other important genes regulated by σ^G include *spoVT,* itself a regulator of σ^G dependent expression (Bagyan *et al.,* 1996). The *gerA* and *gerB* operons which are involved in the germination response to alanine and other germinants are regulated by σ^G along with genes which protect the spore from DNA damage (Hilbert and Piggot, 2004).

1.8.8 Role of σ^K

The final major sigma factor to be involved in *B.subtilis* regulation of sporulation is the mother cell specific σ^K. This sigma factor is expressed specifically in the mother cell via a σ^E dependent promoter which also requires SpoIIID. SpoIIID is a DNA binding protein which activates the transcription of σ^K with σ^E RNA polymerase and σ^K RNA polymerase. σ^k, therefore regulates itself in a positive feedback loop as is the case with σ^G

(Halberg and Kroos, 1994; Kroos *et al.,* 1989). σ^K, like σ^E is expressed in an inactive form which has to be further processed to become active. The pro-σ^K is activated by intra membrane proteolysis in a reaction which is catalysed by SpoIVFB. This protein is under the control of σ^E, and is inhibited by the σ^E dependent BofA. This inhibition is mediated through interaction with SpoIVFA. SpoIVFA is itself inhibited by spoIVB which in turn is inhibited by BofC. Both *spoVB* and *bofC* are regulated by σ^G.

σ^K has been found to regulate genes involved in formation of the spore coat, including *cotC* and *cotD*. The former is repressed by SpoIIID (Wang *et al.,* 2007; Ichikawa *et al.,* 1999; Henriques *et al.,* 1998). σ^K also regulates *spoVD* and *spoVK* which are required for spore maturation (Daniel *et al.,* 1994; Fan *et al.,* 1992). The *gerE* gene is also positively regulated by σ^K (Eichenberger *et al.,* 2004; Steil *et al.,* 2005). GerE is a DNA binding protein which activates certain *cot* genes along with σ^K , including *cotC* and *cotD* (Ichikawa *et al.,* 1999; Ichikawa and Kroos, 2000; Zheng *et al.,* 1992). GerE also positively regulates at least 25 other transcription units in the σ^K regulon and negatively regulates at least 36 transcription units including σ^K (Eichenberger *et al.,* 2004; Ichikawa *et al.,* 1999; Zheng *et al.,* 1992; Wang *et al.,* 2007).

1.9 Sigma factors in *C.acetobutylicum*

Spore formation in *C.acetobutylicum* seems to show many similarities to the process in *B.subtilis*. However, there is still little known about many of the sporulation factors involved. A single transcriptional unit has been annotated as σ^K but this did not exceed the cut off expression criteria stipulated in a recent microarray study (Jones *et al.,* 2008). This study also found the putative *spoIVFB* homologue did not exceed the cut off expression criteria. The expression of two genes (*yabC* and *spsF*) was found to be regulated by σ^K in *B.subtilis*. Both *yabC* and *spsF,* were found to be up-regulated in the mid stationary phase and late stationary phase, respectively. This indicated that *spsF* could be regulated in a σ^K dependent manner. The lack of genetic homologues of *B.subtilis* σ^K

regulated genes has hampered efforts to determine if a σ^K homologue exists in *C.acetobutylicum*. In general this study found the *Bacillus* model to hold true with σ^H followed by Spo0A, σ^F, σ^E and σ^G in order of expression.

Sporulation in the genus of *Clostridium* as a whole is an important developmental process both medically and in the industrial process of solvent production. In *C.acetobutylicum* the ability to reduce the level of sporulation in industrial fermentations would enable higher yields of solvents. More research is needed into the genetic regulation of sporulation in this industrially important organism to achieve this aim.

1.11 Quorum Sensing

Bacteria have evolved many sophisticated mechanisms for cell-to-cell communication, which, in response to environmental cues, results in the coordinated modulation of gene expression (Chan *et al.*, 2004; Lyon and Muir, 2003; Miller and Bassler, 2001; Williams *et al.*, 2000; Winzer *et al.*, 2002).

Quorum sensing can be described as the co-ordinate control of gene expression within a population of bacteria which is cell density dependent (Cámara *et al*., 2002; Williams 2007a; Williams *et al*., 2007b).. This involves the use of intercellular communication signals to activate genes, such as those encoding virulence factors and those which are part of global regulatory systems. Quorum sensing provides the intercellular signals needed to activate a global regulatory system in some bacteria and is achieved through the binding of the signal molecule to a cognate sensor kinase or response regulator protein. This in turn enables a co-ordinate control of gene expression within a given population (Chan *et al.*, 2004).

1.12 Quorum Sensing in Gram-negative bacteria

In Gram-negative bacteria, a typical example of a quorum sensing mechanism in action is the regulation of the luciferase-encoding genes in the bioluminescent marine bacterium *Vibrio fischeri.* When these bacteria are cultured in broth, they exhibit a lag in luminescence gene (*lux*) expression during early and mid-exponential growth followed by a rapid increase in expression in late exponential and early stationary phases. Luminescence was found to be induced in mid log phase by the addition of cell free liquid extracts from stationary phase cultures. The extracts also showed strain specificity as the addition of *Vibrio harveyi* extract to the early log phase of *V.fischeri* (or the addition of *V.fischeri* extract to *Vibrio harveyi* in early log phase) did not induce gene expression at this point. The auto inducer involved was found to be *N*-3-(oxohexanoyl)homoserine lactone (3-oxo-C_6-HSL) (Fuqua *et al.*, 1994; Nealson *et al.*, 1970). This molecule was found to be synthesised by the LuxI protein. The molecule is membrane diffusible and at a threshold concentration, 3-oxo-C_6-HSL activates the intracellular LuxR transcriptional activator which, in the case of *V.fischeri*, results in the transcription of genes responsible for bioluminescence. The Lux I/R system is found to lie under a more complex regulatory circuit involving AinS, *N*-octanoyl HSL (C8-HSL) and LuxS in *Vibrio* species, as seen in figure 1.7.

The *N*-acyl-L-homoserine lactones (AHLs) in Gram-negative bacteria are the most extensively employed sensing signal molecules. They can vary in the length of the acyl chain (between 4 and 14 carbons), saturation levels and the oxidation states of the C-3. The AHLs are usually, but not exclusively, synthesised by members of the LuxI family (Chan *et al.*, 2004).

luxI/R orthologuous genes have been found in a wide group of phylogenetically diverse bacteria including many human pathogens and in many cases two or more *luxI/R* systems will function in a single species (McNab and Lamont, 2003). An example of a human pathogen which

employs multiple LuxI/R homologues is *Pseudomonas aeruginosa.* This species is known to possess two known LuxR/I homologues termed LasRI and RhlRI. Each system modulates a regulon comprising of an overlapping set of genes and the systems interact with LasR activating the transcription of *rhlR* (Latifi et al., 1996; Pesci *et al*., 1997). A third LuxR homologue termed has also been discovered which does not have a cognate acyl-HSL synthase (Chugani *et al*., 2001). This has been known to respond to a variety of AHL's; one of these is the product of LasI (Lee et al., 2006; Oinuma and Greenberg, 2011). A second distinct type of signalling molecule is also produced by *P.aeruginosa*, termed the Pseudomonas Quinolone Signal (PQS) which is a 2-alkyl-4-quinolone (Pesci *et al*., 1999). PQS has been shown to control the production of multiple virulence factors (Deziel *et al*., 2004; Diggle *et al*., 2007; Gallagher *et al*., 2002; Pesci *et al*., 1999). The PQS biosynthetic precursor, 2-heptyl-4-quinoline (HHQ) has also been reported to function as a signalling molecule in *P.aeruginosa* (Xiao *et al*., 2006). With continuing research, it is now becoming evident that complex interaction is occurring between different quorum sensing systems. The discovery of the LuxS system originally in *V.harveyi,* which generates a second signalling system based on an unrelated autoinducer termed AI-2 (see figure 1.7) which is species non-specific (Bassler *et al.*, 1993), raised the possibility of interspecies communication in bacterial communities.

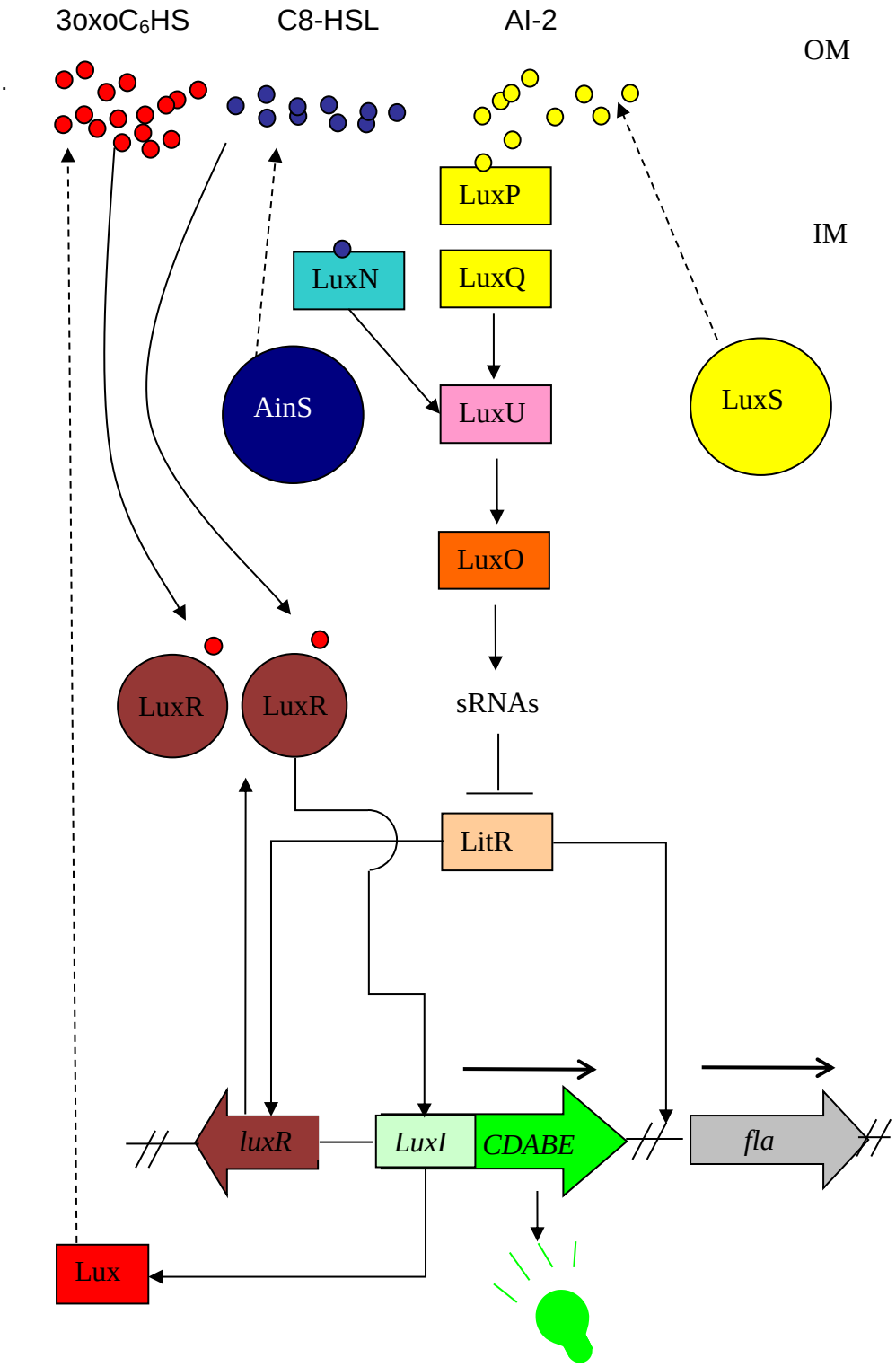


Figure 1.7 The *V.fischeri lux* paradigm regulatory cascade as described by von Bodman *et al* (2008). Some aspects of the model are based on the experimentally defined *V.harveyi* model and remain to be experimentally defined in *V.fischeri*. Three signals are produced by LuxI (3-oxo-C_6-HSL), AinS (C8-HSL) and LuxS (AI-2) represented by red, blue and yellow circles respectively. The latter two signals have associated histidine kinases annotated as LuxN and LuxP/Q, respectively. LuxN and LuxQ undergo autophosphorylation in the absence of signal input, which leads to the phosphorylation of the *LuxO* response regulator. Phosphorylation of LuxO represses *litR* via small RNA (sRNA) negative regulatory elements. The input of signal molecules causes LuxN and LuxQ to perform as phosphatatases which dephosphorylate LuxO through LuxU leaving LitR available to activate several functions including *lux*R and motility genes.

1.13 Quorum Sensing in Gram-positive bacteria

In Gram-positive bacteria, small peptides are most often used as quorum sensing signalling molecules. These are referred to as peptide pheromones (Dunny and Leonard, 1997; Kleerebezem *et al.*, 1997; Nes *et al.*, 1996) or auto-inducing peptides (AIP's). Examples of these include the sporulation stimulating factor (CSF) used for the control of sporulation and genetic competence in *B.subtilis,* which is expressed in early stationary phase (Lazazzera *et al.*, 1997; Solomon *et al.*, 1995). The *phr* system is involved in the sporulation pathway of *B.subtilis*. This organism produces several Phr signals which act as regulators of receptor aspartyl phosphatases (RAPs). Through this interaction they can regulate sporulation by interacting with Spo0F. It has been hypothesised that this phosphorelay system can act as a computational machine performing a sensitive division operation of kinase-encoded signals, in effect sensing the amount of food available per cell which decides whether the cell will enter sporulation (Bischofs *et al.,* 2009).

Other examples of signalling molecules are the γ-butyrolactones. These molecules play a role in the regulation of antibiotic production and sporulation in *Streptomyces* species (Takano *et al.*, 2000).

Two component signalling systems are often involved in quorum sensing in gram positive bacteria. They generally consist of two protein components, a sensor and a response regulator. The sensor proteins involved are usually histidine protein kinases (HPK) which can both autophosphorylate themselves and act as phosphotransferases. They can, therefore, phosphorylate their cognate response regulators (RR) (Wanner, 1992). The sensor proteins usually contain two domains, a variable receptor domain, which detects extracellular or intracellular stimuli, and a conserved transmitter domain where the kinase activity is located. Response regulators are generally also found to consist of two domains. The first is the receiver domain while the second is a highly variable

regulator domain which can bind DNA directly or act with other factors to elicit a response (Wanner, 1992).

The interaction of an external stimulus with the receptor domain of the sensor kinase leads to the autophosphorylation of a conserved histidine residue within the protein's transmitter domain. This activates the phosphotransferase activity of the protein. A phosphoryl group is then transferred to a conserved aspartate residue within the receiver domain of the response regulator. This in turn activates the protein, which via its regulator domain is able to trigger a physiological response, generally by enhancing or inhibiting the expression of specific genes.

1.14 The *agr* System of *Staphylococcus aureus*

In Gram-positive bacteria the most extensively studied quorum sensing system is the accessory gene regulator (*agr*) system of *S.aureus* described in figure 1.8. It is responsible for the co-ordinate regulation of several cell surface proteins and secreted factors produced by the organism (Morfeldt *et al.*, 1988; Peng *et al.*, 1988; Recsei *et al.*, 1986). These secreted factors are up regulated and some cell surface proteins are down regulated towards the end of the exponential phase of growth (Bjorklind and Arvidson, 1980; Lebeau *et al.*, 1994; Vandenesch *et al.*, 1991). The *agr* locus consists of a operon named P2 and a transcript encoding an effector molecule termed P3.

The P2 operon encoding RNA II consists of four genes *agrA, agrB, agrC* and *agrD* (Janzon and Arvidson, 1990; Novick *et al.*, 1995). These four genes are all required for the activation of transcription from the P2 and P3 promoters while the P3 transcript, RNAIII, is itself the effector for the *agr* response (Chan et al., 2004). The products of *agrB* and *agrD* are both needed for the production of AIP (Ji *et al.*, 1995; Ji *et al.*, 1997). *agrD* encodes the prepeptide molecule which is the precursor of the AIP. The prepeptide is processed by the product of the *agrB* gene. AgrB appears to be involved in the processing of AgrD into an octapeptide and its

secretion. It is also involved in the formation of a cyclic thiolactone bond between an internal cysteine and the carboxyl terminus of the peptide (Zhang *et al*., 2002; Yarwood and Schlievert, 2003; Zhang *et al*., 2004; Qiu *et al*., 2005; Kavanaugh *et al*., 2007).

The product of *agrC* is a member of the class 10 histidine protein kinase family (Grebe and Stock, 1999). The autoinducer peptide (AIP) binds to AgrC and may autophosphorylate a conserved histidine residue at its C-terminal domain (Ji et al., 1995; Lina et al., 1998; Lyon et al., 2002). This then results in phospotransfer to AgrA which is required for the activation of P2 and P3 (Ji *et al.*, 1995; Lyon *et al.*, 2002; Novick *et al.*, 1995). The protein AgrA was originally not thought to interact directly with the RNA III promoter region (P3) but to interact with staphylococcal accessory gene regulator protein (SarA) which is the product of a separate genetic locus to mediate activation (Morfeldt *et al.*, 1996a; Morfeldt et al., 1996b). Recent findings have suggested that AgrA binds directly to the P2 and P3 promoters. It has also been shown to bind with a higher affinity to the P2 promoter than to the P3 promoter (Koenig *et al.*, 2004). This raises questions regarding the differential regulation of these two promoters and suggests that the *agr* operon is activated before RNAIII. Indeed the activation of RNAIII may not occur until a certain concentration of AgrA within the cell has been reached (Koenig et al., 2004; Novick *et al.*, 1995). RT-PCR and promoter fusion studies using single, double, triple and complemented mutants have shown that *agrA* is necessary to activate the P2 and P3 promoters (Reyes et al., 2011). The data from this study also indicated that SarA may help activate the P2 promoter by bending the promoter DNA to bring together the AgrA dimers which in turn facilitates the engagement of RNA polymerase to initiate transcription.

Activation of the P3 promoter directs the transcription of RNAIII, which is the effector molecule of the *agr* locus (Peng *et al.*, 1988; Recsei *et al.*, 1986).The production of RNAIII is highest during post exponential and stationary phases of growth and is thought to direct the production of virulence factors at the transcriptional level. Recently it has been shown that there are two distinct subsets of *agr* target gene regulation. This

includes an RNAIII-independent mechanism that is primarily in charge of controlling metabolic and PSM cytolysin genes, and an RNAIII-dependent mechanism which controls virulence genes. The *agr* dependent regulation of the PSM gene family was found to be achieved by direct binding of the AgrA response regulator protein (Queck *et al*., 2008).

The role of *agr* in staphylococcal virulence is complex. The quorum response seems to be highly dependent on the environment in which the organism is grown and there is some evidence, through the use of *agr* repression (particularly RNAIII) in serum and *in vivo* using sub genomic DNA micro arrays, that additional regulators which are not dependent on cell density may also influence *agr* (Yarwood *et al.*, 2002).

There is also evidence that the *agr* phenotype may influence the behaviour and pathogenesis of biofilm-associated *S.aureus* (Yarwood and Schlievert, 2003; Yarwood *et al.*, 2004; Boles and Horswill, 2008; Lauderdale *et al*., 2009; Beenken *et al*., 2010; Kiedrowski *et al*., 2011; Thoendel *et al*., 2011).

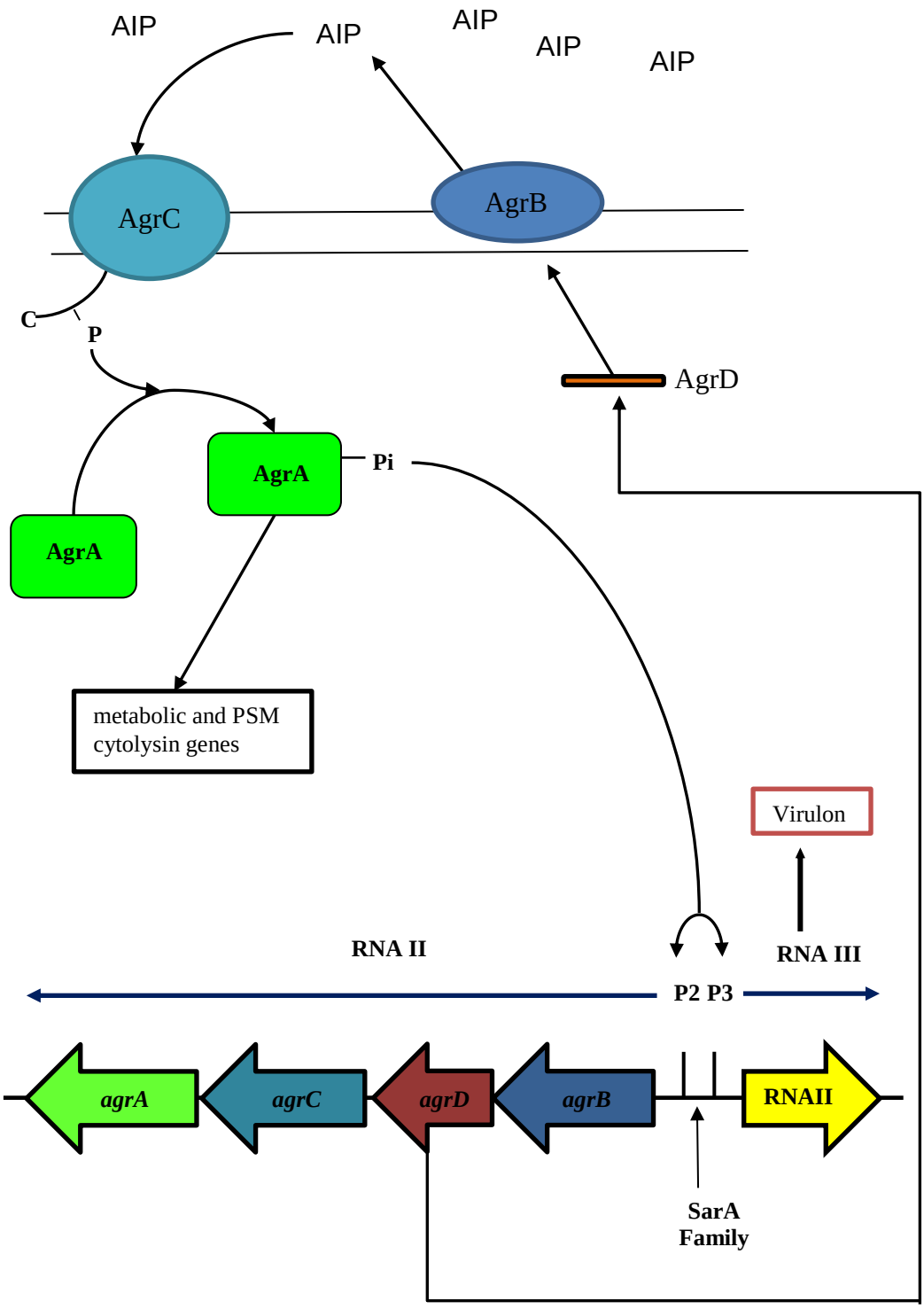


Figure 1.8 The *agr* locus and signalling system. This is a doubly autocatalytic circuit with an activated AgrA inducing more transcription of the other *agr* products. AgrA indirectly has a positive effect on the transcription of RNA III (possibly through interaction with SarA) which itself causes downstream activation and repression of virulence associated factors.

1.15 The role of *agr* like systems in other species

Homologues of the *agr* locus have been found in a variety of Gram-positive bacteria. The *agr* homologues found in Listeria species are found in the same order as in *S.aureus*. There was found to be a high level of conservation in the listerial *agr* genes as with *S.aureus*, which have currently been found to generate four different classes of AIP (Wright *et al*., 2005). No homologue of RNAIII or *sarA* was found in the genome using a BLAST search which suggests the regulation of the *Listeria monocytogenes agr* locus and *agr*-dependent protein expression is different to that of *S.aureus* (Autret *et al.*, 2003).

The role of *agr* in the virulence of *L. monocytogenes* has not been elucidated. Mutants of *agrA* have had varying effects in murine virulence models. An increase in LD50 was found with a Tn*1545* insertion mutant of *agrA* during infection of Swiss mice (Autret et al., 2003). An inframe deletion of *agrA* did not, however affect virulence in a BALB/C mouse infection model which suggests *agrA* does not play a role the regulation of in virulence gene expression (Williams *et al*., 2005). There were some Internalin A-dependent differences in invasion when in vitro assays were performed with exponential phase inocula from a *agrD* deletion mutant. When stationary phase inocula were used, no significant differences were found which suggests a growth dependent effect. An effect on virulence was also seen with bioluminescent imaging when exponential *agrD* mutant cells were injected intravenously (Riedel *et al*., 2009).

Both *agrA* and *agrD* mutations affect the ability of *L.monocytogenes* to adhere to glass and other materials (Garmyn *et al*., 2009; Rieu *et al*., 2007).Suggesting that the *agr* system my affect biofilm formation. Both *agrA* and *agrD* mutants were also found to be affected during the growth as a biofilm under static conditions on a microtitre plate assay in the first 24 hours (Rieu *et al*., 2007).This was repeated under dynamic conditions in a flow cell model of growth (Rieu *et al*., 2008).

In silico analysis of *Lactobacillus plantarum* WCFSI has found two *agr* like two component systems. The first, *lamBDCA* (*Lactobacillus agr*-like

module) contained all four homologues to the *agr* system The second lacked a *lamB* gene and the *agrD* homologue showed some conservation with respect to the *lamD* gene product, but it is not clear whether it is the substrate for a genuine AIP (Sturme *et al*., 2007). The *lamBDCA* operon has the same gene organisation as the *S.aureus agrBDCA* and is transcribed as a single mRNA transcript in a cell density dependent manner. The most abundant peptide purified from the culture media of a LamBD-overexpressing strain, which was designated LamD558, was found to be a cyclopentapeptide thiolactone (Kleerebezem *et al*., 2003; Sturme *et al*., 2002).

Microarray analysis using a *lamA* knockout mutant showed that the *cps2* operon which is one of the exopolysaccharide related operons was upregulated more than 10-fold in this mutant. Phenotypic analysis shows that the *lamA* mutant cells have reduced adherence to glass surfaces. These results suggested that the *lamBDCA* operon may be involved in the cell surface functionality of *L. plantarum* (Fujii et al., 2008).

In *Enterococcus faecalis* three *agr* homologues termed *fsr* have been found which regulate the expression of pathogenicity-related extracellular proteases, *gelE*, a gelatinase gene, and *sprE*, a serine protease gene via a quorum-sensing mechanism (Nakayama *et al*., 2001; Qin *et al*., 2001; Qin *et al*., 2000), and recent studies have suggested that it also regulates biofilm formation (Hancock and Penego 2004; Pillai et al., 2004) and other genes important for virulence (Bourgogne et al., 2006). *fsrA*, *fsrB* and *sprE* mutants in a mouse peritonitis model showed highly significant prolonged survival compared to the parent strain (Qin *et al*., 2000).

In *B.subtilis* the ComX–ComP–ComA signalling pathway regulates the development of genetic competence (Grossman,1995; Lazazzera et al ., 1999; Dubnau and Lovett, 2002) (figure 1.9). The *comX* gene encodes a 10-amino acid peptide that is secreted and accumulates extracellularly (Magnuson et al., 1994; Tortosa et al ., 2001; Ansaldi et al ., 2002). Production of the active ComX pheromone also requires ComQ (Magnuson et al., 1994; Bacon Schneider et al ., 2002). ComX stimulates

the activity of the membrane bound receptor histidine kinase ComP. ComP autophosphorylates and donates phosphate to the response regulator ComA (Comella and Grossman 2005). CSF (also known as PhrC) is a 5-amino-acid peptide that contributes to the activation of ComA by inhibiting the activity of the regulator RapC (Solomon et al., 1995; Core and Perego, 2003). Once phosphorylated and active, ComA functions as a transcriptional activator of its target genes. DNA microarrays have been used to identify genes controlled by the ComX–ComP–ComA quorum-sensing pathway. ComX, ComP and ComA affect the same set of genes, indicating that the kinase ComP is the only receptor for the signalling molecule ComX, and that ComA is the only transcription factor activated directly by ComP, under the conditions tested (Comella and Grossman 2005). The same study also found the expression of over 20 genes appears to be controlled directly by this signalling pathway, and expression of over 150 additional genes, including those involved in competence development, appears to be controlled indirectly.

An *agr*-type quorum-sensing system has also been found in microarray analysis of *Roseburia inulinivorans* a common anaerobic polysaccharide-utilizing firmicute bacterium from the human colon This system was found to be up-regulated during growth on fucose.

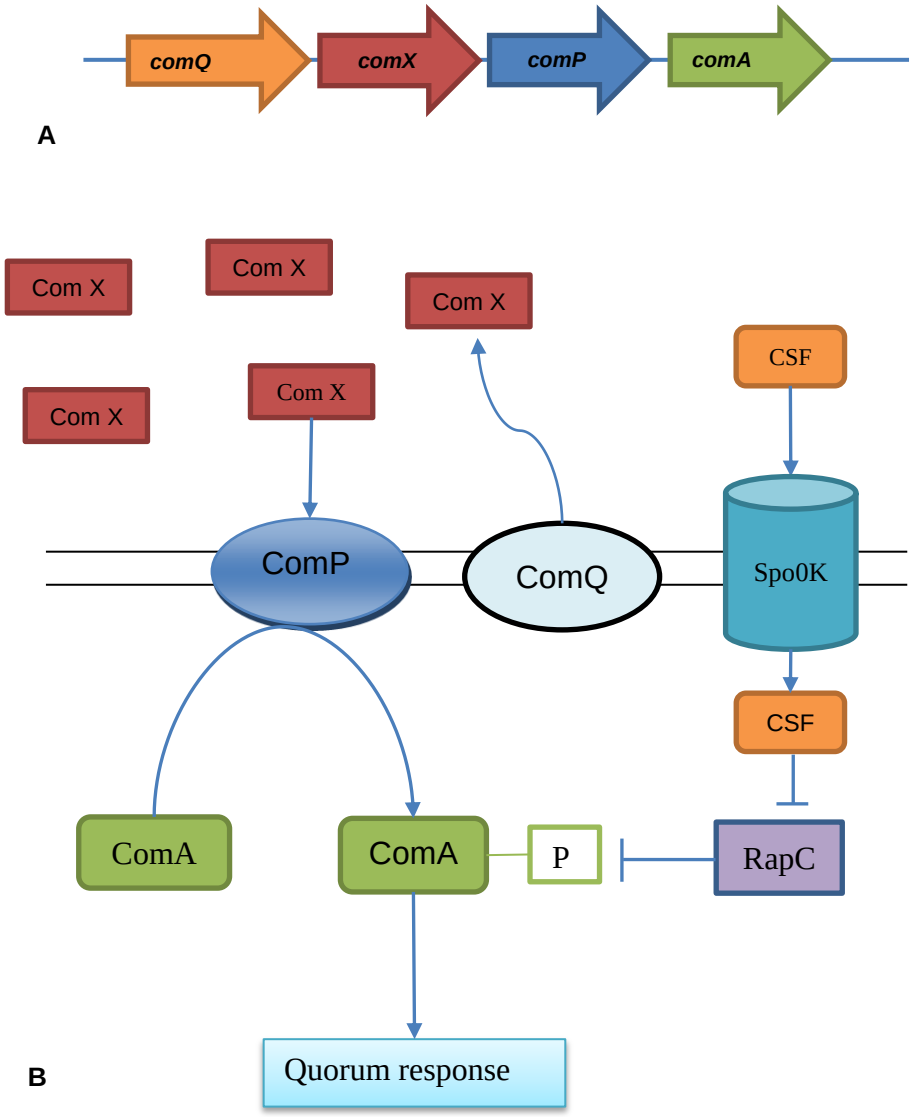


Figure 1.9 (A) The organisation of the *com* gene cluster. (B) The pathways involved in an *agr* like quorum sensing response of *B.subtilis*. The pheromones responsible for regulating the activation of this response are ComX and the CSF (also known as PhrC). ComX accumulates extracellularly and activates the histidine kinase ComP. ComP undergoes autophosphorylation and donates the phosphate to ComA, leading to the activation of ComA-dependent promoters. The CSF pheromone also stimulates the Com regulated response by inhibiting RapC which inhibits ComA activity.

1.16 Quorum sensing in *Clostridium*

The genus *Clostridium*, one of the largest prokaryotic groups, is considered of high industrial and medical importance. The clostridia demonstrate an extraordinary ability to thrive in specialised environments and adapt to a wide range of different stresses. These stress factors often provide the niche needed for the species to out compete other flora, as with gut colonisation by *Clostridium difficile* following antibiotic therapy. It

would be reasonable to think that such adaptive responses require co-ordinate control of gene expression within a population and a cell-cell signalling molecule(s) is most likely involved.

Several studies have shown that clostridia elaborate signal molecules. *Clostridium perfringens* is a causative agent of many human and animal diseases including gas gangrene (Hatheway, 1990; Rood and Cole, 1991). Many extracellular toxins are produced by *C.perfringens* including α-toxin (phospholipase C), θ-toxin (perfingolysin O), κ-toxin (collagenase) and sialidase. Studies have revealed a two component regulatory system consisting of the VirR and VirS proteins which has been shown to be involved in the regulation of these toxins (Lyristis *et al.*, 1994; Shimizu *et al.*, 1994). There have been reports of a diffusible signalling substance which appeared to be involved in the regulation of toxin production (Higashi *et al.*, 1973; Imagawa *et al.*, 1981; Tatsuki *et al.*, 1981). This signalling molecule 'substance A' may have a direct effect on the VirR/VirS system. In mutational studies using two types of theta toxin mutant, one type was found to produce substance A and the other was deficient in this molecule. It was found that the mutant producing substance A could restore the theta toxin activity in the mutant which was deficient in producing this molecule. This activity of substance A appeared in exponentially growing cultures and disappeared in stationary phase (Imagawa and Higashi, 1992).

The locus homologous to both *agrB* and *agrD* was identified using an EZ-Tn*5*-based random mutagenesis approach (Vidal *et al* 2009). This mutant was found to be deficient in perfringolysin O and alpha toxin production. The *agrB* and *agrD* genes were found to be co-transcribed in an operon with two upstream genes encoding hypothetical proteins. The region encoding the *agrB* and *agrD* genes was deleted in a further study (Ohtani *et al* 2009). Pefringolysin O was again found to be deficient in this strain and there were significantly reduced levels of α- and κ-toxins. This effect was complemented by the addition of wild type culture supernatant suggesting the presence of a signalling peptide. A double mutant of the

agr and *virR*/*virS* system could not be complemented with this supernatant indicating that VirR/VirS may be involved in the sensing of this putative substance. These two studies indicate that the *agr* system has a major role in the virulence of this organism.

The identity of substance A has yet to be elucidated. Spent supernatants of *C. difficile* have also been found to contain a small molecular weight molecule which is capable of stimulating toxin A production (Roberts and Minton, unpublished data).

1.16.1 The *agr* homologue in *C.acetobutylicum*

The switch from acidogenesis to solvent production in *Clostridium saccharoperbutylacetonicum* strain N1-4, a close relative of *C.acetobutylicum* is also thought to be influenced by a small diffusible molecule (Kosaka *et al*., 2007). The annotated *C.acetobutylicum* genome contains genes encoding homologues of *agrA*, *agrB* and *agrC*. The existence of a gene encoding AgrD had not been noted. However, a closer inspection of the region immediately adjacent to *agrB* has revealed the presence of a small unannotated ORF capable of encoding a small protein of 54 amino acids. This protein sequence shares some homology with the AgrD peptide of *S.aureus.* Most significantly it contains the Cys residue required for the post-translational cyclisation-mediated production of an AIP.

The above observation suggests that *C. acetobutylicum* contains an equivalent locus to the *S. aureus agr* operon. Furthermore, its discovery has allowed the identification of equivalent loci in other clostridial species. Thus, a BLASTP search of the *C. perfringens* genome with AgrD identified a small ORF encoding a 45 amino acid protein which retains the 'Cys' residue in the C-terminus. Furthermore this ORF is preceded by CPE1561 which encodes a polypeptide identified by BLASTP as containing a conserved AgrB like domain. Similar genes are also evident in the recently completed *C. difficile* and C*lostridium botulinum* genomes, with AgrD

peptides of 49 and 45 amino acids in size, respectively. Unlike *C. acetobutylicum* however, *C. perfringens*, *C. difficile* and *C. botulinum* do not possess *agrCA* two-component systems adjacent to the identified *agrBD* loci.

The discovery of a potential quorum sensing system in clostridia represents a significant discovery. *C. perfringens* is a major animal pathogen. *C.botulinum* is a potentially very serious food pathogen and on the CDC list of potential biological weapons. However, from a human health perspective, *C.difficile* represents perhaps the most important clostridial species. This organism is known for its resistance to antibiotics and is a major cause of healthcare associated disease. It has been notoriously difficult to deal with long term antibiotic induced infections and alternative therapies to antibiotics have been actively sought. Along with the other clostridia it lives up to its name in being difficult to work with.

1.17 Genetic manipulation of *Clostridium*

The discovery of an *agr* system in *C.acetobutylicum* raises many questions as to what role it may play. The two main areas of study with this organism are into its product formation and sporulation. It would be interesting to see if the *agr* system had a role in these two metabolic and physiological events. The lack of genetic tools available for the study of *Clostridium* has impeded much of the research into this organism.

1.17.1 Generation of random mutants

The development of random mutational tools has played a major role in assigning a gene's function. The model tools used for this are transposons such as Tn*5* and Tn*10*. One frequently employed transposon in *Enterococcus faecalis* is Tn*916* which shares homology with conjugative elements of *C. difficile* (Hachler *et al.,* 1987). As a result this transposon has been used to try and achieve mutational insertion in a variety of

different clostridia. Initial work with the *C.difficile* strain CD37 resulted in the transposon inserting into a “hot spot” (Mullany *et al.,* 1991; Wang *et al.,* 2000) but it was found to insert randomly in other *C.difficile* strains (Roberts *et al.,* 2003; Hussain *et al.,* 2005; Sebaiha *et al.,* 2005).

Research into transposon mutagenesis in *C. perfringens* using Tn*916* also revealed this tendency to insert in multiple copies. Analysis of mutants containing a single copy of Tn*916* which were deficient in the production of virulence factors identified the gene encoding *virS*, part of the two component *virRS* system which plays an important role in the regulation of virulence in *C. perfringens* (Lyristis *et al.,* 1994; Banu *et al.,* 2000).

The use of Tn*916* as a mutagenic tool has also been employed with solventogenic clostridia. The conjugative transfer of Tn*916* into *C. acetobutylicum* DSM 792 from *E. faecalis* has been demonstrated but only half of the transconjugants contained a single copy of the transposon (Bertram and Durre, 1989).

When Tn*916* was introduced into *C.acetobutylicum* ATCC 824, over half of transconjugants obtained were found to have multiple copies inserted into the same site (Mattsson and Rogers, 1994). Various classes of mutants were isolated which were defective in the production of solvents and/or sporulation. Mutants involved in solvent formation and sporulation have also been isolated in *Clostridium saccharobutylicum* P262 using Tn*916*. This led to the identification of the *sum* gene which is thought to be essential for clostridial differentiation and sporulation in this species (Collet *et al.,* 1997).

The use of transposons for generating mutants has had limited use partially due to the fact they have a tendency to insert into a site in multiple copies and into so called “hot spots”. With the publication of the whole genome sequences of many clostridial species, efforts have been made to target specific genes of interest using site directed methods.

1.17.2 Directed gene inactivation using host mediated recombination

With the recent publication of a variety of genomic sequences the search for methods of targeted gene knockout has gained momentum. Directed inactivation through homologous recombination has proved a successful method of assigning gene function in a variety of different bacteria and eukaryotes. Such strategies employ the cell's own recombinational machinery to introduce an extrachromosomal element such as a plasmid into the targeted gene. The extrachromosomal element will contain an inactivated copy of the gene. A rare recombination event between this inactivated gene in the plasmid and the host gene can then occur.

Single crossover recombination (figure 1.10) involves one recombinational event resulting in the entire element inserting into the recipient's genome via a Campbell-like mechanism. If the single crossover event is to be mutagenic then the element ideally has to insert into the central portion of the target gene. This is therefore an insertional form of gene inactivation as with transposons. Upon integration the two copies in the host's targeted gene lack either a 5′ end or a 3′ end of the gene. The resulting two copies are, therefore, non-functional.

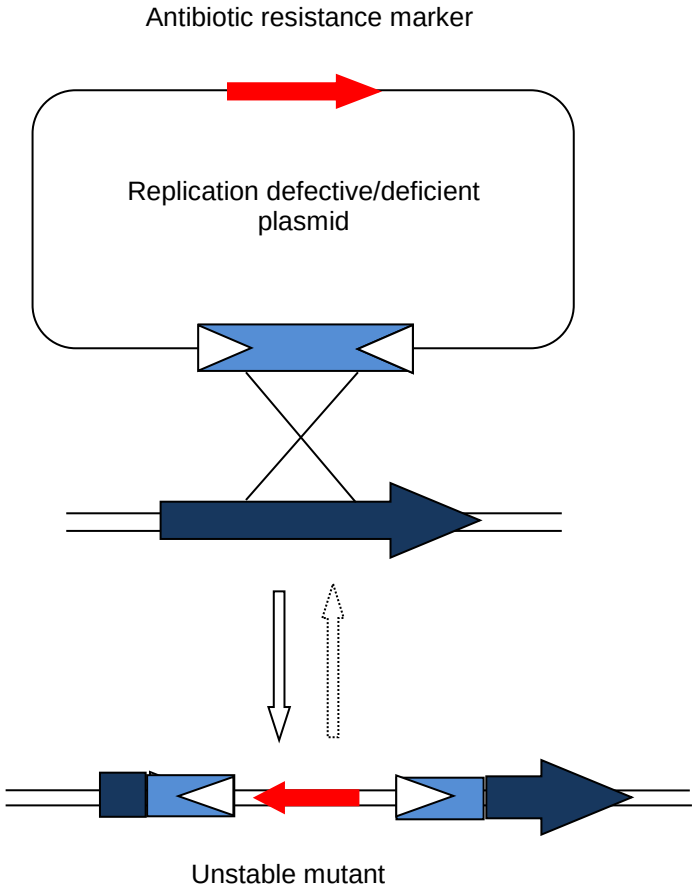


Figure 1.10 A single crossover event leading to an insertional mutation. In this case the element introduced into the host is a plasmid which is usually replication defective/deficient. The plasmid also contains an antibiotic marker allowing selection of integrants. This kind of mutation is unstable as a second recombinational event can occur which leads to excision of the whole plasmid and reversion to the wildtype genotype.

The recombinational events necessary for mutation are rare so mechanisms to allow selection for these events are desirable. For this reason plasmids which are defective or deficient in replication are utilised. The most common plasmids employed are termed suicide vectors. These plasmids often carry an antibiotic marker on the plasmid backbone which can be selected for when the plasmid has integrated rather than being lost from the cell.

A double crossover event occurs across two regions of homology where the intervening segment of DNA is inserted into the genome and replaces the target host gene by reciprocal exchange (figure 1.11). The fragment

that replaces the host gene therefore needs to be inactivated either through insertion or deletion of DNA. The inserted DNA can be a selective marker such as an antibiotic resistance gene. A second antibiotic marker can be placed on the vector backbone to reveal if the vector has been lost from the host.

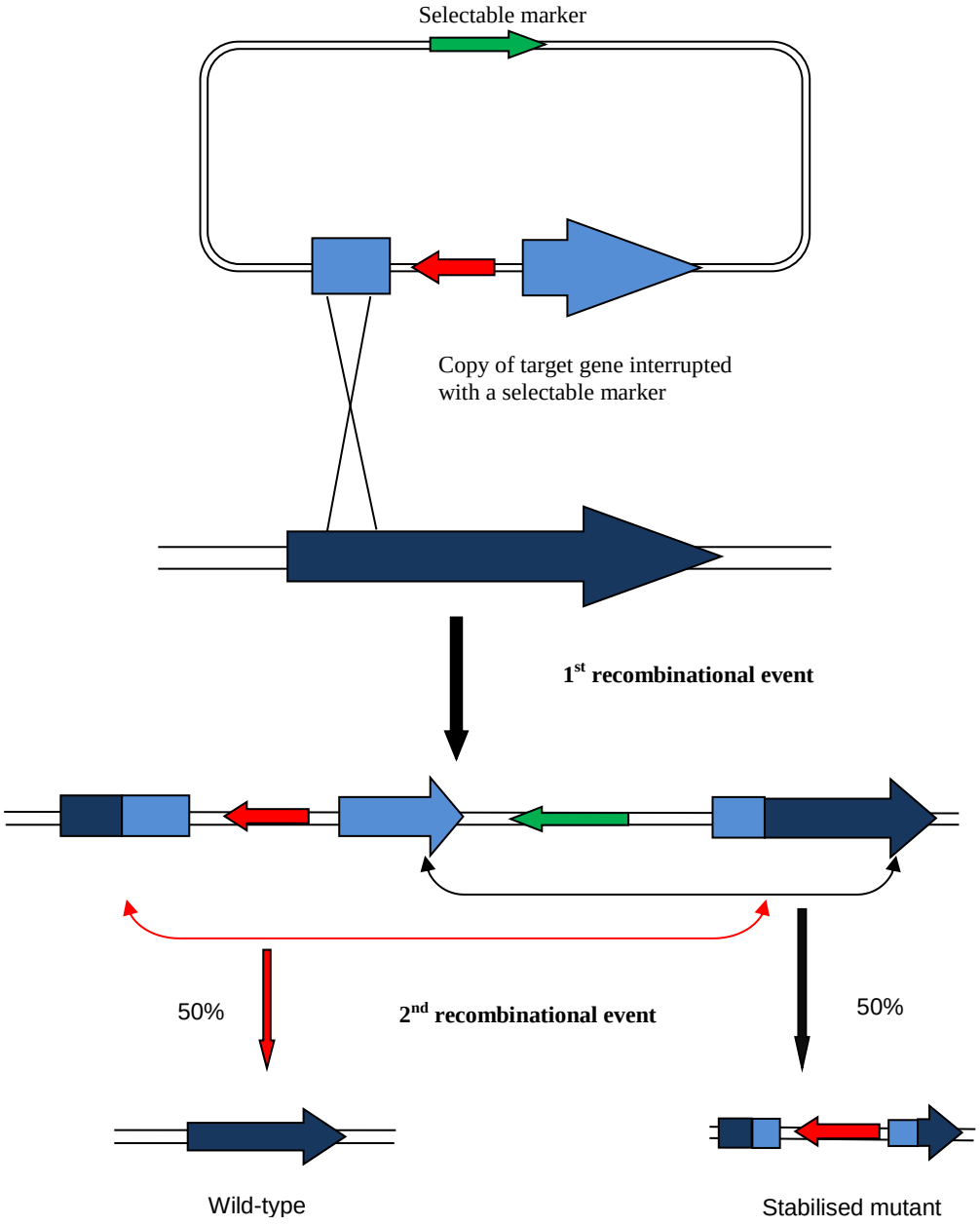


Figure 1.11 Double crossover event in which two recombinational events can lead to a stabilised mutant. As with a single crossover event an unstable plasmid is often used. This can contain a copy of the gene to be targeted which is interrupted with a selectable marker. After the first recombinational event the whole plasmid is inserted into the genome but still leaves the target gene intact. The second recombinational event can lead to a full excision of the plasmid leading to a reversion to the wild-type genotype. The second recombinational event can alternatively lead to portions of the host gene and plasmid DNA excising which will lead to a stable mutant. The selectable markers within the truncated homologous target gene and the plasmid backbone allow the latter event to be selected for.

Both single and double crossover targeted gene mutagenesis have been used to obtain gene knockout in the genus *Clostridium*. The first reported examples were achieved in *C.perfringens*, in particular strain 13 (Shimizu *et al*., 1994; Awad *et al.,* 1995). The ability to obtain mutants in this particular strain is probably due to the fact that it is particularly amenable to DNA transfer by electroporation. This may be due to the absence of a type II restriction system (Shimizu *et al.,* 2002). A number of mutants have been generated in this strain using suicide plasmids specifically designed for a single crossover event. Genes targeted include *pfoA* (Awad *et al.,* 2001), *pfoR* (Awad and Rood, 2002), *luxS* and *virX* (Ohtani *et al.,* 2002). Double crossover mutants have also been obtained using sequential generation of single crossover mutants (O'Brien and Melville, 2004; Varga *et al.,* 2006).

The first reported examples used replication deficient suicide vectors which contained either a *tet* (Shimizu *et al.,* 1994) or *ermBP* (Awad *et al.,* 1995) marker inserted into a plasmid based copy of the target gene. These plasmids had no selective marker on the plasmid backbone. It was only possible to distinguish between those which had undergone a single recombinational event and the double crossover mutant using a phenotypic plate test. The use of two antibiotic genes was first employed to target the *cpe* gene in the disease relevant strains SM101 and FM4969 (Sarker *et al.,* 1994) and was subsequently used to target *colA* in strain 13 (Awad *et al.,* 2000).

Directed gene inactivation has also been used to target genes in *C.difficile.* The first gene to be targeted in this species was *gld*A (Liyanage *et al.,* 2001). They used a method previously employed to obtain a gene knockout in *C.beijerinckii* (Wilkinson and Young, 1994), where an *oriT* based suicide plasmid pMTL31 was conjugated into a non-toxigenic strain CD37. Internal fragments of the *gldA* gene were cloned into this plasmid which also contained the *ermB* antibiotic resistance marker. Integrants resistant to erythromycin were selected, but disruption of *gldA* appears to

be lethal and resulted in pin point colonies. However, PCR analysis did reveal that the plasmid had integrated into the gene.

A successful recombination-based gene knockout system has been described for the genomic strain of *C.difficile* (CD630) which was cured of its *ermB* gene. The plasmid used, pJIR1456 was based on pIP404 which has been found to be deficient in replication in *C.difficile*. This plasmid was used to target 2 genes *rgaR* and *rgbR* through single crossover integration. A looped out, independently replicating plasmid was found in some of the cells indicating that these mutants were unstable. The *codY* gene was also targeted in this strain by using the suicide plasmid pSD21 carrying a *catP* marker (Dineen *et al.,* 2007). This mutant was found only to be stable in the presence of thiamphenicol. Without this selection the culture became predominantly wild type, due to the excision of the integrated plasmid.

Due to the potential economic and environmental benefits it is not surprising that there have been attempts to achieve single insertion and double crossover targeted gene knockout in solventogenic species. As previously stated directed gene insertion was first achieved in *C.beijerinckii* NCIMB 8052 by targeting the genes *gutD* and *spo0A* (Wilkinson and Young, 1994). These mutants were obtained by the conjugative transfer of pMTL30, a *oriT*-based suicide vector in which internal fragments of the two genes were inserted. The same method was used to achieve the knockout of *scrB* and *scrR* in a later study (Reid *et al.,* 1999). Unusually, multiple insertions of the plasmid were found when the mutants were characterised.

With *C.acetobutylicum* ATCC 824, five published mutations have been reported. Four of these involved a single crossover insertion of a replication deficient plasmid. The inactivated genes were *butK*, *pta*, *aad* and *solR*, the latter two genes residing on pSOL-1 (Green *et al.,* 1996; Green and Bennett, 1996; Nair *et al.,* 1999). With these mutants the

suicide plasmids used the target site flanked by 2 directly repeated copies of the DNA segment directing integration. As a result they were segregationally unstable.

The gene *spo0A* was targeted using a method designed to try and achieve a double crossover mutation (Harris *et al.,* 2002). This method was similar to the strategy reported by Sarker *et al.* (1999). It involved the use of a plasmid which incorporated the pIM13 replicon commonly used in cloning studies in *C.acetobutylicum*. This plasmid pETSPO contained the *spo0A* gene interrupted with *ermB* and another selective marker *catP* on its backbone. Those cells transformed with the plasmid were transferred onto reinforced clostridial medium containing erythromycin every 24 hours for 5 days. Colonies were then replica plated onto media containing erythromycin or thiamphenicol to select those which may have undergone a double crossover event. A single colony with the desired characteristic was obtained. Unfortunately it was subsequently found that a true double crossover event did not take place. Rather, a crossover event occurred between two homologous sequences which were present at the 3′ end of the *repL* structural gene and upstream of the *ermB* gene. Loss of the 3kb fragment between these 10 nt sequences resulted in the inactivation of the *spo0A* gene through insertion of a 2.1 kb fragment containing *ermB*.

A similar study was reported at Clostridium IX (Rice University, Houston, TX USA, 18-21 May 2006) in which a pIM13 based plasmid was used to generate a double crossover knockout of the genes encoding the Cac824I type II restriction enzyme and *ctfAB* in pSOL-1 (Saint-Prix,2006). The knockout cassette was constructed using the flanking regions of the target gene interrupted with the *ermB* selective marker. The selective marker was therefore in effect replacing the target gene. Furthermore the *ermB* marker was flanked by FRT sites. This enabled the *ermB* to be excised by the addition of the yeast FLP recombinase gene on a second plasmid.
The resulting mutants were, therefore, less prone to polar effects and the absence of a marker allows the possibility for subsequent knockout of other genes in the strain.

1.17.3 Antisense RNA

Antisense RNA has been used as a natural mechanism of gene control in all three kingdoms and has been used successfully by molecular biologists to down regulate genes in a variety of organisms (Rasmussen *et al.,* 2007).

In *C.acetobutylicum* the gene for glutamine synthase (*glnA*) is controlled by antisense RNA. This gene has a promoter P3 located downstream which controls a transcript of RNA which is complementary to the start of the *glnA* mRNA. This antisense RNA is thought to regulate *glnA* in a nitrogen dependent manner (Fierro-Monti *et al.,* 1992).

The use of this technology relies on the production of a single stranded mRNA molecule which is complementary or “antisense” to the mRNA of the target gene. The molecule can, therefore, bind to the target mRNA molecule through base pairing with it. The resultant double stranded molecule is unable to be translated, and, therefore, the target gene product is not produced. It is thought this inhibition is caused by either one of two mechanisms. The inhibition of the ribosome binding to the mRNA or the increased degradation of the double stranded RNA through the cell’s natural mechanisms.

This method of gene silencing has been successfully used in a variety of clostridial species (Raju *et al.,* 2007; Perret *et al.,* 2004; Liyanage *et al.,* 2000; Liyanage *et al.,* 2001; Nakayama *et al.,* 2008). Most notably this method has been used to silence genes in *C.acetobutylicum* (Tummala *et al.,* 2003a; Tummala *et al.,* 2003b; Tummala et al., 2003c; Scotcher *et al.,* 2005; Desai *et al.,* 1999).

1.17.4 Clostron

Group II introns can be described as ribozymes in reference to the fact they are capable of catalysing their own splicing out of a primary RNA transcript. They are found in both eukaryotic organelles and prokaryotic mRNA. Their excision does not require GTP which distinguishes them from group I introns.

The spliced out RNA forms a lariat structure which is highly similar to the lariats formed by nuclear introns of eukaryotes. It is thought that these may have originally evolved from group II introns.

Some group II introns can be mobile and insert themselves into an intronless copy of the host gene. These introns contain an orf encoding a protein termed the 'intron encoded protein' (IEP). They use a mechanism by which the excised intron RNA reverse splices directly into a DNA target site and is then reverse transcribed by the IEP (Lambowitz and Zimmerley 2004). This mechanism is termed retrohoming and is an efficient process occurring in very high frequencies in both organelles and bacteria (Lazowska *et al.,* 1994; Moran *et al.,* 1995; Cousineau *et al.,* 1998; Cousineau *et al.,* 2000).

Alan Lambowitz and colleagues at the University of Texas have undertaken a detailed study of the mechanistic details of the replication process of the mobile group II intron from the Ll.ltrB gene of *Lactococus lactis*. They determined the rules which allow target site recognition and showed that by following these rules they could re-target the Ll.ltrB intron into original sites of the *E.coli thyA* gene *in vivo*.

An algorithm was developed which predicted the optimal insertion sites for this intron and designed appropriate primers for modifying the intron to insert within these sites. This made it possible to target almost any gene of interest, limited only by the minimal recognition sequence determined by the IEP. Furthermore generating the retargeted introns by PCR without cloning made it theoretically possible to achieve gene disruption in a very short period of time in both Gram-negative and Gram-positive bacteria

where the LI.ltrB intron can function (Perutka *et al.,* 2004). These intron's which are targeted towards an original gene are referred to as Targetrons. A further development was to design Targetron vectors which have the IEP ORF *in trans* of the intron's. The IEP ORF can be removed from within the intron without affecting the introns mobility. The Orf deletion donor intron no longer encodes the IEP which itself is lost with the plasmid. The intron cannot, therefore, splice out and further mobility is prevented. This averts the possibility of secondary mutations (Karberg *et al.,* 2001).

The host range of Targetrons has proved to be extremely valuable in a wide variety of bacteria (Karberg *et al.,* 2001; Frazier *et al.,* 2003; Yao *et al.,* 2006). It has also been used to target the alpha toxin (*plc*) gene in *C.perfringens* (Chen *et al.,* 2005). The frequencies of integration vary widely between target sites which could make screening with PCR laborious if a low frequency was achieved or if a phenotypic screen was unavailable.

A unique solution to this problem was devised by Zhong and co-workers (Zhong *et al.,* 2003). Domain IV of the intron sequence which normally encodes the IEP is non-structural. This makes possible the insertion of a cargo sequence within this domain. They inserted an antibiotic marker in reverse orientation under the control of its own promoter into domain IV. This marker was itself disrupted by an efficiently self-splicing group I intron in the forward orientation. During retrotransposition of the Targetron via an RNA intermediate the group I intron was spliced out. This activated the selectable marker which is then integrated into the target DNA and allows selection of integrants using an appropriate antibiotic. This type of marker was described as a Retrotranscription – Activated selectable Marker (RAM) and was found to enable one-step bacterial gene interruption at near 100% efficiency (figure 1.12).

The Lambowitz laboratory has patented this technology through a spinoff company called Ingex which licensed the technology to Sigma Aldrich.

1.18 Development of the clostridial RAM

There were obvious benefits of using a Targetron based system to achieve gene knockout in clostridia. For this reason it was decided to create a new RAM. The main body of this work was undertaken by John Heap (Heap *et al.,* 2007). The RAM system used by Zhong *et al.* (2003) used an ITPG inducible T7 promoter and was based on a trimethoprim resistance gene. The Sigma Aldrich derivative was very similar but used a kanamycin selective marker. The T7 promoter does not work in clostridia and clostridial species are naturally resistant to both these antibiotics. For this reason the *ermB* gene from *Enterococcus faecalis* was chosen as a selectable marker and the group I intron employed was that of the *td* gene of phage T4.

The RAM designated ErmbtdRAM2 includes the td intron and its exons in a linker inserted between the *ermB* ORF and its promoter which prevents erythromycin resistance. The splicing out of this intron from the reverse strand yields a modified *ermB* gene which encodes a functional protein with 12 additional amino acids at its N-terminus. The RAM is controlled by a thiolase promoter (Girbal *et al.,* 2003) from *C.acetobutylicum* as the original *ermB* promoter was deemed too weak to confer resistance from a single chromosomal copy (figure 1.12). This RAM was incorporated into the IV domain of the Ll.ltrB intron in pMTL007. The intron in pMTL007 is itself under the control of the IPTG inducible *fac* promoter. The plasmid also contains the chloramphenicol resistance gene *catP* within its backbone which allows selection of the plasmid in both *E.coli* and clostridial species (figure 1.12).

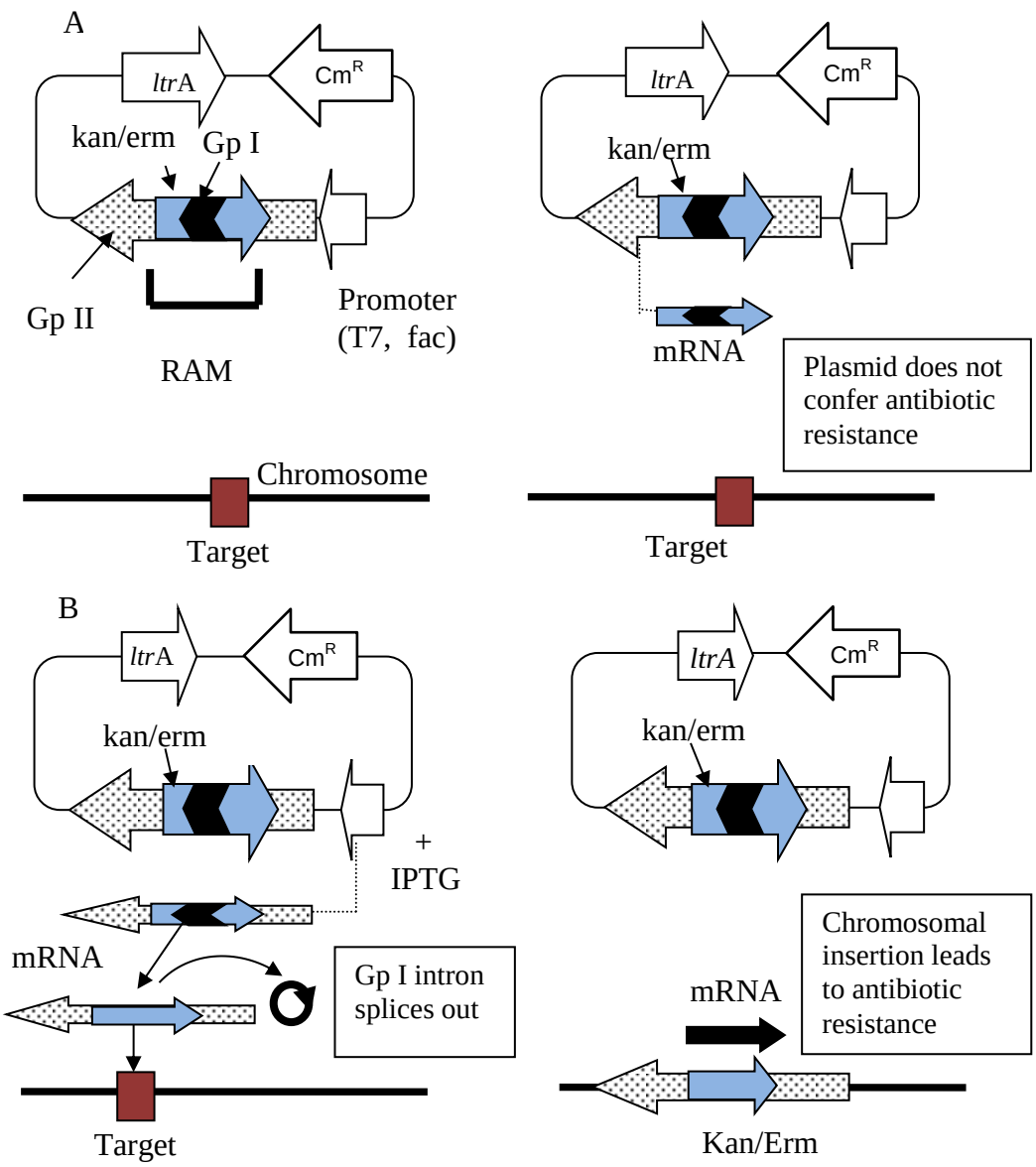


Figure 1.12 The process of insertional deletion using the Targetron and Clostron systems.
(A) The vector containing the targeting sequence is transformed into the cell. At this stage the group I (GpI) intron prevents transcription of the antibiotic resistance gene (erythromycin (erm) or kanamycin (kan)) which rests inside the group II (GpII) intron. The vector also has the *ltrA* gene which encodes the protein machinery for the targeted insertion of the group II intron.
(B) Expression of the group II intron with or without IPTG induction causes the group I intron to splice out of the RAM marker in the messenger RNA. The group II intron can then insert into the target sequence with an intact antibiotic marker.

Much of this study has been involved with developing genetic tools for the study of *agr*. *In C.acetobutylicum* the ClosTron system is a recently developed tool which has been used to obtain gene disruption and will aid the analysis of the whole *agr* system. By knocking out the *agr* system it will be possible to look at the phenotypic consequences for the organism. As this has not been achieved before and *C.acetobutylicum* is a non-pathogenic organism and therefore does not produce an assayable toxin, the obvious choice is to look for a possible effect on product formation and sporulation.

AIMS

C.acetobutylicum is an industrially important organism which could have great potential in producing valuable chemicals and potentially an economically important biofuel. The discovery of a putative *agr* system is of great interest as it plays a major role in the regulation of genes in other bacterial species. The first aim of this study was to take advantage of the latest methods to obtain gene disruption of the *agr* genes in *C.acetobutylicum*. This involved optimising methods of transformation of the species and developing useful and reproducible methods for obtaining gene disruption.

Once gene disruption was achieved the second aim was to identify any physiological changes which may occur in *C.acetobutylicum agr* knockouts. The level of product formation, including acids and solvents, was measured along with the levels of sporulation. The *agr* system in *C.acetobutylicum* may be involved in a global regulatory role in *C.acetobutylicum* and may play a role in quorum sensing. For this reason microarray technology was used to assess the effect of *agrB* knockout on the transcriptome.

2.0 MATERIALS AND METHODS

2.1 Bacterial Strains and Plasmids

Table 2.1 Bacterial strains and plasmids

Organism/ plasmid	Genotype/ features	Source/ reference
Strain		
Clostridium acetobutylicum ATCC 824	Genomic strain	P.Soucaille Toulouse
Escherichia coli Top10	F$^-$ *mcr*A Δ(*mrr-hsdRMS-mcrBC*) φ80*lacZ*ΔM15 Δ*lacX*74 *deoR recA*1 *araD139* Δ(*ara-leu*)7697 *galU galK rpsL*(*Str*^{R}) *endA1 nupG*	Invitrogen
Escherichia coli Top10F'	F'{*lacI*^{q} Tn*10* (Tet^R)} *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lacZ*ΔM15 Δ*lac*X74 *deoR recA1 araD139* Δ(*ara-leu*)7697 *galU galK rpsL(Str^R) endA1 nupG*	Invitrogen
B. subtilis BS34A	Tcr *B subtilis*::Tn*916* 11	A.Roberts; UCL
Clostridium difficile CD630	Genome strain PCR type 12 (pathogenic)	Hachler *et al*., 1987
Plasmid		
pAN1	The ϕ3T I methyltransferase gene of *B. subtilis* phage ϕ3tI inserted into the *tet* gene of pACYC184 Cm^R	Mermelstein and Papoutsakis 1993a
pAN2	The ϕ3T I methyltransferase gene of pAN-1 Tet^R	S.Cartman Nottingham University
pCR2.1-TOPO	E. coli PCR cloning vector, ColE1, Ap^R, Km^R	Invitrogen

pCR2.1::Cac*agrB*	*E. coli* PCR cloning vector, ColE1, Ap^R, Km^R containing *agrB* promoter fragment	This study
pGC4lacIgusAoriT	*E.coli* Clostridium Gus expression vector *lac*I, *tra*J, Cm^R	G.Carter Nottingham University
pGC4lacIcatPoriT	*E.coli/Clostridium* CAT expression vector *lacI, traJ,* Cm^R	G.Carter Nottingham University
pCD6ColE1catPoriT	*Ecoli/Clostridium shuttle vector traJ,* Cm^R	G.Carter Nottingham University
pIMP1	*Ecoli/Clostridium* shuttle vector Em^R	Mermelstein *et al*., 1992
pSOS95	*Ecoli/Clostridium* shuttle vector *thl* promoter, Amp^r MLS^r ColE1 Ori repL	Souicaille, P., and E. T. Papoutsakis. Unpublished data.
pMTL9361	IPTG inducible clostridial expression vector. Em^R, pCD6 replicon, P_{ptb}::*lacI*, P_{fac}, *lacZ' catP*	G.Carter Nottingham
pJSARD916	IPTG inducible clostridial expression vector. Em^R, pCD6 replicon, *B.subtilis*::Tn*916*	This study
pJSARD5397	IPTG inducible clostridial expression vector. Em^R, pCD6 replicon, *C.difficile::Tn5392*	This study
pJS01	pGC4 based *E.coli Clostridium* Gus expression vector with *agrB* promoter Cm^R	This study
pJS02	*E.coli Clostridium CAT* expression vector with *agrBD* promoter Cm^R	This study
pJS03	*E.coli/Clostridium* CAT expression vector with *agrC*A promoter Cm^R	This study
pJS04	pIMP1 based *E.coli/ Clostridium* CAT expression vector with *agrB*D promoter Cm^R	This study

pJS05	pIMP1 based *E.coli/ Clostridium* CAT expression vector with *agrC*A promoter Cm^R	This study
pSOS95	*E. coli/Clostridium* shuttle vector carrying the *thl* promoter	Tummala *et al*., 2003
pJS06	*E. coli/Clostridium* shuttle vector carrying antisense *agrB* short strand under control of the *thl* promoter	This study
pJS07	*E. coli/Clostridium* shuttle vector carrying antisense *agrC* short strand under control of the *thl* promoter	This study
pJS08	*E. coli/Clostridium* shuttle vector carrying antisense *agrB l*ong strand under control of the *thl* promoter	This study
pJS09	*E. coli/Clostridium* shuttle vector carrying antisense *agrC* long strand under control of the *thl* promoter	This study
pCONS2.1	Mobilisable *Clostridium* vector, Cm^R	C.Croux Toulouse
pREPbuk 3	Mobilisable *Clostridium* Knockout vector targeted to *buk*, contains FRT MLS insert Cm^R Em^R	C.Croux Toulouse
pJS10	Mobilisable *Clostridium* Knockout vector targeted to *agrA*, contains FRT MLS insert Cm^R Em^R	This study
pCLF-1	Mobilisable *Clostridium* vector, contains FLP recombinase under thiolase promoter,Cm^R	C.Croux Toulouse
pMTL007	Clostron vector contains group II intron, ErmBtdRAM2 and *ltrA* ORF, Cm^R	J Heap *et al*., 2007
pMTL007Ca-*agrA*-141a	pMTL007 based vector targeted to *agrA*, Cm^R	This study
pMTL007Ca-*agrB*-385a	pMTL007 based vector targeted to *agrB*, Cm^R	This study

pMTL007Ca-*agrC*-121s	pMTL007 based vector targeted to *agrC*, Cm^R	This study
pMTL007C-E2	Clostron vector contains group II intron, ErmBtdRAM2 flanked by FRT target sites and *ltrA* ORF, Cm^R	John Heap Nottingham
pMTL007C-E2::Cac-*agrA*-141a	Clostron vector based on pMTL007C-E2 targeted to *agrA*	This study
pMTL007C-E2::Cac-*agrB*-385a	Clostron vector based on pMTL007C-E2 targeted to *agrB*	This study
pMTL007C-E2::Cac-*agrC*-121s	Clostron vector based on pMTL007C-E2 targeted to *agrC*	This study
pMTL960	*E.coli/Clostridium* shuttle vector with thiolase promoter Cm^R	N. P. Minton Nottingham
pJS11	pMTL960 based complementation vector with *agrC* and *agrA* genes	This Study
pJS12	pMTL960 based complementation vector with *agrB* and *agrD* genes	This Study

2.2 Chemical Reagents

All general chemicals were obtained from Sigma Aldrich unless otherwise stated.

2.3 Growth Media

2.3.1 Luria-Bertani medium

Luria-Bertani (LB) broth was prepared as described by Sambrook *et al.* (1989) and consisted of 10 g tryptone (Difco Laboratories), 5 g yeast extract (Difco Laboratories) and 5 g sodium chloride made up to 1 L with distilled water. Sterilisation was achieved by autoclaving at 120 °C and 15 psi for 20 min.

LB agar was prepared by addition of 1% (w/v) Number 1 Bacteriological Agar (Oxoid).

2.3.2 2 x YT Medium

2 x YT broth was prepared as described by Sambrook *et al.* (1989) and consisted of 16 g tryptone (Difco Laboratories), 10 g yeast extract (Difco Laboratories) and 5 g sodium chloride made up to 1 L with distilled water.

2 x YT agar was prepared by addition of 1% (w/v) Number 1 Bacteriological Agar (Oxoid).

2.3.3 2 x YTG Medium

2xYTG medium was prepared as 2xYT with the pH adjusted to 5.2 using HCl before making it up to 900ml and autoclaving. Once cool, 100ml of sterile 20% glucose was added giving a 2% final glucose concentration.

2.3.4 Clostridial Basal Medium CBM

This medium was prepared based on a recipe described by O'Brien and Morris (1971) and consisted of 200 mg magnesium sulphate (heptahydrate), 10 mg manganese sulphate, 10 mg iron sulphate

(heptahydrate), 1 mg p-aminobenzoic acid, 2 µg biotin, 1 mg thiamine.HCl, 4g casein hydrolysate (Oxoid) in 800 ml of distilled water. For CBM agar 10 g of Number 1 Bacteriological Agar (Oxoid) was added at this stage. Sterilisation was achieved by autoclaving at 120 °C and 15 psi for 20 min. The medium was then made up to 1 L by the addition of 50 ml 20% glucose solution w/v, 10 ml of 50 mg ml^{-1} potassium phosphate monohydrate, 10 ml of 50 mg ml^{-1} potassium phosphate dihydrate and sterile distilled water.

2.3.5 Reinforced Clostridial Medium RCM

This medium is based on the formula used for Oxoid Reinforced Clostridial Medium (RCM) but does not contain agar. 3 g of yeast extract, 10 g of Lab-Lemco powder, 10 g of peptone, 5 g of glucose, 1 g of soluble starch, 5 g of sodium chloride, 3 g of sodium acetate and 0.5 g of cysteine hydrochloride were made up to 1 litre with distilled water. Sterilisation was achieved by autoclaving at 120 °C and 15 psi for 20 min.

2.3.6 Clostridial Basal Sporulation Medium CBSM

Clostridial Basal Sporulation Medium (CBSM) consisted of 10 ml of CBM salts (8 g magnesium sulphate (heptahydrate), 0.4 g iron sulphate (heptahydrate) , 0.3 g manganese sulphate in 400 ml of distilled water), 4 g enzymatic casein hydrolysate, 20 ml glucose (20% (w/v)) and 15 g Bacteriological agar No 1 (Oxoid) in 930 ml of distilled water, 12.3 g magnesium sulphate (heptahydrate), Sterilisation was achieved by autoclaving at 120 °C and 15 psi for 20 min. The medium was then made up to 1 L by addition of 10 ml of filter sterilised CBM vitamins (10 mg p-aminobenzoic acid, 10 mg thiamine HCl and 20mg biotin in 100 ml of distilled water) 20 ml of filter sterilised potassium phosphate buffer (80 ml of 1 M dibasic potassium phosphate mixed with 90 ml of 1M monobasic potassium phosphate; adjusted to pH 7.0 with 1 M dibasic potassium phosphate solution and made up to 200 ml with distilled water) and sterile

distilled water. A sterile 250 g l^{-1} solution of calcium chloride was added to the medium to give a final concentration of 5 g l^{-1}.

2.3.7 Synthetic Medium

This medium was made based on a recipe from the Hubert Bahl laboratory in Rostock. It consisted of 60 g glucose 0.55 g potassium phosphate (monobasic), 0.55 g potassium phosphate (dibasic), 0.22 g magnesium suphate heptahydrate, 2.3 ml acetic acid, 0.11 g iron sulphate (heptahydrate). The pH was adjusted to pH 6.6 with ammonium hydroxide and 21.3 g of 2-(*N*-morpholino)ethanesulfonic acid (MES),8 mg of para-aminobenzoic acid (PABA) and 0.08 mg of biotin was added before the medium was made up to 1 L and sterilised by autoclaving at 120 °C and 15 psi for 20 min.

2.4 Supplements

Growth media were supplemented at the following concentrations as required: ampicillin (Amp), 50 $\mu g.ml^{-1}$; chloramphenicol (Cm), 25 $\mu g.ml^{-1}$; thiamphenicol (Thm), 10 $\mu g.ml^{-1}$, tetracycline (Tet), 10 $\mu g.ml^{-1}$; erythromycin (Erm), 500 $\mu g.ml^{-1}$ (for clostridia 2.5 $\mu g.ml^{-1}$); kanamycin (Km), 200 $\mu g.ml^{-1;}$ isopropyl-1-thio-β-D-galactopyranoside (IPTG) 10µM. Stock solutions were prepared as described by Sambrook *et al*. (1989) and stored according to the manufacturer's instructions

2.5 Growth Conditions

E. coli was grown aerobically at 37 °C unless otherwise stated. Liquid cultures were grown in Luria Bertani (LB) broth or YTG broth with gentle agitation at 200 rpm. Strains of *E. coli* were maintained in frozen bead stocks at -80 °C.

All clostridia were grown at 37 °C in an anaerobic cabinet (MK3 Anaerobic Work Station, Don Whitley Scientific) containing an atmosphere of 80%

(v/v) nitrogen, 10% (v/v) hydrogen and 10% (v/v) carbon dioxide, unless otherwise stated.

CBM medium was used for the general propagation of *C. acetobutylicum*, 2 x YTG medium supplemented with 0.5% (w/v) glucose was used for transformation of *C.acetobutylicum*. Strains of clostridia were maintained in CBSM stored at -20°C or in CBM with 20% glycerol at -80°C. Sporulation assays, product analysis and the microarray were undertaken with *C.acetobutylicum* grown in synthetic medium.

Growth of all bacterial cultures in liquid medium was monitored by optical density at 600 nm (OD_{600}) using a Pharmacia Novaspec II.

2.6 DNA Manipulations

2.6.1 Plasmid Preparation

Plasmid preparation was carried out by the alkaline lysis method described by Sambrook *et al*. (1989) using Qiagen spin miniprep kits (supplied by Qiagen Ltd. UK) as described in the manufacturer's instructions.

2.6.2 Restriction Digests

Restriction enzymes were obtained from New England Biolabs, Roche Pharmaceuticals or Promega. Digests were carried out as described in the manufacturer's instructions. Briefly, digests contained 1 x buffer supplied by manufacturer, 0.5 – 1 μg DNA, 5 – 20 U restriction endonuclease, unless the endonuclease exhibits star activity when 1 U was added. The appropriate final volume was made up with water and maintained with < 5% v/v glycerol concentration. DNA was cleaned before further use as described in section 2.6.7.

2.6.3 Blunt-Ending of DNA fragments Using Klenow Polymerase

DNA with 5' overhanging ends was blunt-ended using DNA Polymerase I, Large (Klenow) Fragment, obtained from New England Biolabs or Promega as described in the manufacturer's instructions. Briefly, to a restriction digest (described above), 1x DNA polymerase buffer (supplied by manufacturer), 2.5 U Klenow polymerase and 33 μM each dNTP was added, followed by incubation at 25 °C for 15 min.

2.6.4 Blunt-Ending of DNA Fragments Using T4 DNA Polymerase

DNA with 3' overhanging ends was blunt-ended using T4 DNA Polymerase, obtained from New England Biolabs, as described in the manufacturer's instructions. Briefly, to a restriction digest (described above), 1 x T4 DNA polymerase buffer (supplied by manufacturer), 5 U T4 DNA polymerase and 100 μM each dNTP was added, followed by incubation at 12 °C for 15 min. The reaction was cleaned using dialysis as described in 2.9.1.

2.6.5 Dephosphorylation of Linearised DNA Fragments

Linearised DNA fragments were dephosphorylated using Shrimp Alkaline Phosphatase, obtained from Roche Pharmaceuticals, as described in the manufacturer's instructions. Briefly to a restriction digest (described above), 1 x SAP buffer (supplied by manufacturer) and 2 U Shrimp Alkaline Phosphatase was added, followed by incubation at 37°C for 1 h.

2.6.6 Ligation of DNA fragments

Ligation of DNA fragments was carried out using T4 DNA ligase, obtained from Roche Pharmaceuticals or Promega, according to manufacturer's

instructions. Briefly, purified DNA fragment inserts were added to purified vector fragments at ratios of 1:1, 3:1 and 5:1. Ligations were carried out in 10μl volumes also containing 1 x T4 DNA ligase buffer (supplied by manufacturer) and 3 U ligase, followed by incubation at 16 °C overnight.

Blunt-end ligations were additionally supplemented with 10% (w/v) Polyethylene Glycol 8000.

2.6.7 PCR and Restriction Digest Clean Up

DNA fragments from (RT-) PCR reactions, restriction digests and blunt-ending reactions were routinely purified away from contaminating oligonucleotides and enzymes using Qiagen's PCR Purification Kit as described in the manufacturer's instructions.

2.6.8 Extraction of DNA Fragments from Agarose

DNA fragments were visualised using ethidium bromide and the desired band excised from the agarose gel using a clean scalpel. The DNA was then extracted from the excised agarose slice using Qiagen's Gel Extraction Kit as described in the manufacturer's instructions.

2.7 Agarose Gel Electrophoresis

Agarose gels were prepared and run as described by Sambrook *et al*. (1989). Gels were prepared in 1 x TAE (40 mM Tris Acetate, 2 mM EDTA, pH8.5) at a concentration of 1- 2% (w/v) agarose. Ethidium bromide was added to a final concentration of 10 $\mu g.ml^{-1}$. Gels were run in 1x TAE buffer at 50 – 120 V. DNA was visualised using a UV transilluminator.

2.8 Polymerase Chain Reaction

Table 2.2 Oligonucleotide primers used in this study

Primer name	Primer Sequence	Application
TN916F	5´TAGGGATCCTTGAATCTCTACAAAGAA AGGAC-3′	ARD amplification
TN916R	5´-TGTCTCGAGAGATTTAATAGACG ATTTCAA-3′	ARD amplification
TN5397F	5´-ACAGGATCCTTGAACTGCCAGAGT AGGAA-3′	ARD amplification
TN5397R	5´-ATGCTCGAGCTGATCTTCCAGCGTTA AT-3′	ARD amplification
AGRBDF	5´-TCATCTAGATTCATTACAAAC TCCACCTA-3′	Amplifying *agrB* promoter region and RT-PCR analysis of antisense vector targeted to *agrB* and *agrD* *agrB complementation vector*
AGRBDR	5´-CAAGAGCTCAACTTCTGCTAACTTTT CC-3′	Amplifying *agrB* promoter region
AGRCAF	5´-GCCTCTAGAGCTACTTAAGGAGAATA TTG-3′	Amplifying *agrC* promoter region and RT-PCR analysis of antisense vector targeted to *agrC* and *agrA* *agrC/A* *complementation vector*
AGRCAR	5´-TATGAGCTCATTATAATTCCAACTAA TTC-3′	Amplifying *agrC* promoter region
CDcond1	5´-CCTAATGCATATCTATTTATTAA-3′	primer used for detection of pMT9361 conjugated into *C.acetobutylicum* 824
CDcond4	5´-CCCTAAGCCGACCAAGCAGAG-3′	primer used for detection of pMT9361 conjugated into *C.acetobutylicum* 825
GUTDF	5´-CCCTACCAGGAGGAATTTTTATGAAA AA-3′	strain characterisation (conjugation)
GUTDR	5´-GCATGCATCTTTTTTTATTTAACCTC TTGTTT-3′	strain characterisation (conjugation)
AGRARVCONF	5´-GAGGACAATCAGTTTAAG-3′	characterisation of whole gene KO in *agrA* agrC/A complementation vector
FprimerBDcat	5´-CCCATCGGCGAGGCTAGTTACCC-3´	CAT reporter vectors
RprimerBDcat	5´-CGGCCAGTGAATTCCCCT-3´	CAT reporter vectors
FprimerCAcat	5´-GCCATGGTACCCATCGGCGAG-3´	CAT reporter vectors
RprimerCAcat	5´-CGACGGCCAGTGAATTCCCC-3´	CAT reporter vectors
agrBDFRT	5´-CAGGCTGTACTTATGGGAATGC-3′	RT-PCR *agr* junctions
agrBDRRT	5´-GTAGCTCTTCAGGCATTTTAGGT-3´	RT-PCR *agr* junctions
agrDCFRT	5´-CGAACCTAAAATGCCTGAAGAGCT AC-3´	RT-PCR *agr* junctions
agrDCRRT	5´-GGCATCCAGCTACAGGTAGAGA-3´	RT-PCR *agr* junctions
agrCAFRT	5´-GAGGATTTTCTACAAAAGGTG-3´	RT-PCR *agr* junctions

agrCARRT	5´-GCAGCTAGCTGTATCCCATTTA TAG-3´	RT-PCR *agr* junctions
agrBAasFBam	5´-GCGGATCCGGTTAAGGTCGTATAATT AA-3′	antisense vector targeted to *agrB* and *agrD*
agrBASRBAM	5´-GCGGATCCAAATGCAGCAAGCATT ATAATAG-3´	antisense vector targeted to *agrB* and *agrD*
agrCasFBam	5´-GCGGATCCTATTTGTCAATATTCTC CTTAAG-3′	antisense vector targeted to *agrC* and *agrA*/ Characterisation of agrA whole deletion
agrCasRBam	5´-GCGGATCCAGTGTATTGAAACATCAT TT-3′	antisense vector targeted to *agrC* and *agrA*
CAASRTP	5´-GCCATAACCGTAACCTGTTGCC-3´	RT-PCR analysis of antisense vector targeted to *agrC* and *agrA*
BDASRTP	5´-CTCTATACCCTCCATTTTTAGATGC-3´	RT-PCR analysis of antisense vector targeted to *agrB* and *agrD*
asBD2GENR04 /1	5´-CAGGATCCCTTGTACCTCCTTCCTA ATACC-3′	second longer antisense vector targeted to *agrC* and *agrA*
asCA2GENR04 /1	5´-CAGGATCCGTAATCATGCCTAAACC CTCTC-3´	second longer antisense vector targeted to *agrB* and *agrD*
BamXhoPR1ex press	5´-GATAGGATCCTCGAGGTAATACTAT ATCTCTTAAATTATAC	*agrB* complementation vector
RVRprimagrC	5´-CAGCAGCTAGCTGTATCCC-3´	*agrC/A* complementation vector
RVNARIagrCA RP	5´-CAGGCGCCGTGAGCCGTCTACTAT TAC	*agrC/A* complementation vector
ERY5'R	5´-AAGGGTAAAATGCCCTTTTCCTGAGC CG-3´	Characterisation of mutants with *ermB* Marker
ERY3'D2	5´-GATAAACAGAAGTATAATTATTTCG TTATGAAATGGG-3´	Characterisation of mutants with *ermB* Marker
agrAKOF1	5´-GGGATCCGTATCAAATATGATGCA G-3′	*agrA* whole deletion mutant
agrAKOF2	5´-GTATGCGAAGATAATAAGGCCTAG GATGTTAAAGGGAC-3′	*agrA* whole deletion mutant
agrAKOR1	5´-GTCCCTTTAACATCCTAGGCCTTAT TATCTTCGCATAC-3´	*agrA* whole deletion mutant
agrAKOR2	5´-CCACGGATCCTATATCTCACCTGAT ATACCCAGGATCCCTTGTACCTCCTTCC TAATACC-3′	*agrA* whole deletion mutant
CONAGRKOR	5´-CACATGATACCCCACATAAACC-3′	Characterisation of *agrA* whole deletion
CACamyPF	5´-CCTTTAGCGCTGTTGTTCC-3′	amplifys *amyP* from pSOL-1 pSOL-1
CACamyPR	5´-GGTGTCTCCAGTTCAAGG-3′	amplifys *amyP* from pSOL-1 pSOL-2
Cac-agrA-141a-IBS	5´-AAAAAAGCTTATAATTATCCTTAATTT TCAATGACGTGCGCCCAGATAGGG TG-3′	ClosTron *agrA* mutagenisis
Cac-agrA-141a-	5´-CAGATTGTACAAATGTGGTGATAACA	ClosTron *agrA*

EBS1d	GATAAGTCAATGACATTAACTTACCTTT CTTTGT-3′	mutagenisis
Cac-agrA-141a-EBS2	5′-TGAACGCAAGTTTCTAATTTCGGTTA AAATCCGATAGAGGAAAGTGTCT-3′	ClosTron *agrA* mutagenisis
Cac-agrC-121s-IBS	5′-AAAAAAGCTTATAATTATCCTTAAAC TGCAATTATGTGCGCCCAGATAGGG TG-3′	ClosTron *agrC* mutagenisis
Cac-agrC-121s-EBS1d	5′-CAGATTGTACAAATGTGGTGATAACA GATAAGTCAATTATTTTAACTTACCTTTC TTTGT-3′	ClosTron *agrC* mutagenisis
Cac-agrC-121s-EBS2	5′-TGAACGCAAGTTTCTAATTTCGGTTC AGTTCCGATAGAGGAAAGTGTCT-3′	ClosTron *agrC* mutagenisis
Cac-agrB-385a-IBS	5′-AAAAAAGCTTATAATTATCCTTACTGT ACCTCCAGGTGCGC-3′	ClosTron *agrB* mutagenisis
Cac-agrB-385a-EBS1d	5′-CAGATTGTACAAATGTGGTGATAAC AGATAAGTCCTCCAGG-3′	ClosTron *agrB* mutagenisis
Cac-agrB-385a-EBS2	5′-TGAACGCAAGTTTCTAATTTCGATTT ACAGTCGATAGAGGA-3′	ClosTron *agrB* mutagenisis
EBS UNIVERSAL	5′-CGAAATTAGAAACTTGCGTTCAGTA AAC-3′	Clostron general
agrAasR	5′-GATATGAGACATTCTTCTCC-3′	Characterisation of Clostron mutants with *ermB* RAM
agrBasF	5′-GGTTAAGGTCGTATAATTAAATTA AAG-3′	Characterisation of Clostron mutants with *ermB* RAM
agrCasF	5′-GAGATATAGTATTACAAAAAAGG GG-3′	Characterisation of Clostron mutants with *ermB* RAM
ErmRam-F	5′-CGAAATTAGAAACTTGCGTTCAGTA AAC-3′	Clostron general
5402F-F1	5′-GCTTGGCTCCAGGTCGACGG-3′	Clostron general
007-R1	5′-GCTGGCGAAAGGGGGATGTG-3′	Clostron general
SEQAGRBPF	5′-ACGCGTTATATTGATAAAAATAATAA TAGTGGG -3′	Characterisation of *agrB* ClosTron Mutants
SEQAGRBPR	5′-TTAAGGAGGTGTATTTCATATGACCA TGATTACG-3′	Characterisation of *agrB* ClosTron Mutants
SEQAGRCPF	5′-AGGGTATCCCCAGTTAGTGTTAAGTC TTGG-3′	Characterisation of *agrC* ClosTron Mutants
SEQAGRCPR	5′-GTTAGCAGAAGTTGTTTCTTTG -3′	Characterisation of *agrC* ClosTron Mutants
SEQAGRAPF	5′-GAGCCTCTGTATCTCCAGGCG -3′	Characterisation of *agrA* ClosTron Mutants
SEQAGRAPR	5′-CGTATCAATTTTTAGCCTG -3′	Characterisation of *agrA* ClosTron Mutants
Cac-agrB-sF1	5′- CCACAGAATTATTAGGAGC-3′	Characterisation of *agrB* ClosTron Mutants
Cac-agrB-sR1	5′-GTATGCGAAGATAATAAGG -3′	Characterisation of *agrB* ClosTron Mutants
Cac-agrC-sF1	5′-CGTCCAATTCACCCTCTATATCC -3′	Characterisation of *agrC* ClosTron Mutants
Cac-agrC-sR1	5′-TATATGGTTCTGGCGTCTTTTATAAT CGTAAACC -3′	Characterisation of *agrC* ClosTron Mutants
Cac-agrA-sF1	5′- TTGTCCCTCAGTTTAGCTCCAACTAA AGG-3′	Characterisation of *agrA* ClosTron Mutants
Cac-agrA-sR1	5′- TATTCATGCAGAATTAAATGGAAAAG GCTTGTAGG-3′	Characterisation of *agrA* ClosTron Mutants

KOVECP1R	5´-GCTCGTCAGGGGGGCGGAGCCTATGG-3´	Characterisation of *agrA* whole knockout Mutant
KOVECP2R	5´-CTATTCTAACTATATCATAACTGTTC-3´	Characterisation of *agrA* whole knockout Mutant

2.8.2 General PCR Parameters

PCR reactions were carried out using the Failsafe PCR system (Cambio) as described in the manufacturer's instructions. Briefly, on ice, 0.5 μl of each primer (from 100 μM stock), 100 ng DNA template, and 1.25 U of Failsafe PCR enzyme mix were mixed and the reaction was made up to a final volume of 25 μl with sterile, nuclease free water. Finally, 25 μl of 2 x Failsafe PCR premix buffer was added bringing the final reaction volume to 50 μl. The mixture was then subjected to PCR in a thermal cycler block.

Failsafe PCR premix E was routinely used for all PCR reactions. If premix E failed to yield a product, then the reaction was repeated with the remaining 11 premixes (A-D & F-L); the one that gave the highest yield was used in subsequent repeat PCRs.

Annealing temperatures varied and were dependent upon the primers used. They were calculated by assuming a temperature of 2 °C for A and T nucleotides and 4 °C for C and G nucleotides. The annealing temperature used in the individual PCRs was 5 °C below the lowest calculated temperature for the primer pair. An example of the parameters based on the Failsafe PCR systems recommendations is as follows:

- Initial denaturation of the template at 98°C for 1-2 min.
- 3-step cycling program, 40 cycles
- Denature at 95°C for 0.5-1 min.
- Annealing of the primers at a temperature 5°C below the lowest calculated annealing temperature for the primer pair for 0.5-1 min.
- Extension of the annealed primers at 72°C for 1 min for every kb of expected product.

After amplification, the samples were kept at 4°C overnight or frozen at –20°C.

2.8.3 Colony PCR

A single colony was taken from an agar plate and resuspended in 10 μl of sterile distilled water. The cell suspension was then heated to 98 °C for 10 min, before centrifugation at 16,100 x g in a bench top centrifuge. A 5 μl aliquot was carefully removed (avoiding any cellular debris) and added to the PCR reaction.

2.9 Transfer of Plasmid DNA into Bacterial Cells

2.9.1 Dialysis of DNA Ligation Reactions

Ligation reactions were dialysed against water using 0.025 μm dialysis filters (Millipore Corporation). The filter was allowed to float, shiny side up, in a petri dish containing sterile deionised water. The ligation reaction was then added to the uppermost surface by pipette and left to dialyse for 30-60 min.

2.9.2 Preparation of Electrocompetent *E. coli*

A 5 ml aliquot from an overnight culture of *E. coli,* grown in LB-broth (with selection if necessary), was used to inoculate 250 ml sterile LB-broth, followed by incubation at 37 °C with shaking (200 rpm), until an OD_{600} of 0.5 - 0.8 was reached. The culture was chilled on ice for 15 – 30 min, before the cells were harvested by centrifugation at 5000 x g for 15 min at 4 °C. The supernatant was aspirated and the pellet resuspended in 250 ml of ice-cold sterile distilled water. The cell suspension was then centrifuged as before, washed in 100 ml of ice-cold sterile distilled water and resuspended in 0.5 ml of 10% (v/v) glycerol. 40 μl aliquots were transferred to chilled microfuge tubes before storage at -80 °C

2.9.3 Electroporation of Plasmid DNA into *E.coli*

Dialysed plasmid DNA or ligation reactions were added to 40 μl of electrocompetent *E. coli* (as prepared in **Section 2.9.2**), that had been allowed to thaw on ice for 10 min. This mixture was then transferred to a pre-chilled 2 mm gap electroporation cuvette (Biorad). A pulse of 2.5 kV (25 μF, 200 Ω) was delivered to the cuvette using a BioRad Gene Pulsar according to the manufacturer's instructions. A 700 μl aliquot of SOC or 2 x YT medium was immediately added, and the cell suspension was incubated at 37 °C for 1 h with shaking (200 rpm). 100 μl, 10 μl and 1 μl aliquots were removed from the cell suspension and if necessary made up to a final volume of 100 μl with pre-warmed 2 x YT medium before being spread onto 2 x YT agar containing the appropriate selection.

2.9.4 Preparation of Chemically Competent *E. coli*

A 1 ml aliquot from a 10 ml overnight culture of *E. coli* (grown with selection if necessary) was used to inoculate 50 ml of sterile LB-broth, which was then incubated at 37 °C with shaking (200 rpm), until an OD_{600} equivalent to 0.4 was achieved. The cells were harvested by centrifugation at 5000 x g for 3 min at 4 °C, before being resuspended in 25 ml of ice cold 0.1 M $MgCl_2$. The cells were harvested as before and the pellet gently resuspended in 2 ml of ice cold 0.1 M $CaCl_2$. The cell suspension was then left on ice for a minimum of 2 h, to achieve chemical competence. 100 μl aliquots could then be used for chemical transformations for up to 48 h provided the cell suspension was kept chilled.

2.9.5 Transformation of Plasmid DNA into Chemically Competent *E. coli*

Plasmid DNA or ligation reactions were added to 100 μl of competent *E. coli* cells that had been kept chilled on ice and gently mixed. The transformation reactions were then incubated for exactly 30 min on ice, followed by heat-shock at precisely 42 °C for 30 sec. The mixture was then incubated on ice for a further minute, before the addition of 200 μl SOC or 2 x YT medium. The reaction was then incubated at 37 °C for one h with shaking, 200 rpm before plating onto 2 x YT agar containing the appropriate selection.

2.9.6 Methylation of plasmid DNA for transfer into *C.acetobutylicum*

Before transfer into *C. acetobutylicum* plasmids were purified from *E.coli* TOP10 containing either pAN-1 (Mermelstein and Papoutsakis 1993) or pAN-2 (Heap *et al.,* 2007) which contain the *B. subtilis* phage ϕ3tI methyltransferase. This protected the plasmids from the *C. acetobutylicum* Cac824I DNA restriction activity.

2.9.7 Conjugation of Plasmid DNA into *C. acetobutylicum*

Plasmids were introduced into *C.acetobutylicum* by conjugation from *E. coli* CA434 essentially as described by Purdy *et al*. *(*2002). A 1 ml aliquot from a 5 ml overnight culture of the donor strain, grown in LB-broth with appropriate selection, was taken and the cells harvested by centrifugation in a bench top microfuge at 5000 x g for 1 min. The pellet was gently resuspended with a pipette tip in 1 ml sterile PBS and the cells harvested as before. The pellet was resuspended in 200 μl aliquots of *C.acetobutylicum* 824 (P.Soucaille, Toulouse) anaerobically, overnight at 37 °C. The cell suspension was then spotted in 10 μl volumes onto non-selective agar plates, followed by incubation at 37 °C for 7 – 8 h under anaerobic conditions. The cells were then harvested by flooding the agar plates with sterile PBS. The resulting cell slurry was removed, and spread

on selective agar plates, followed by incubation at 37 °C for 16 - 96 h in an anaerobic workstation.

2.9.8 Electroporation of Plasmid DNA into *C. acetobutylicum*: using *C.* beijerinckii method

Plasmid DNA was introduced into *C. beijerinckii* NCIMB 8052 essentially as described by Oultram *et al*. (1988a). *C. acetobutylicum* was serially diluted 1:10 into a final volume of 5 ml of 2 x YTG medium ten times and grown anaerobically at 37 °C overnight. The lowest two dilutions that showed growth were then used to inoculate 100 ml sterile 2 x YTG medium, which was then incubated anaerobically at 37 °C until an OD_{600} of 0.6 was achieved. The culture was cooled on ice for 30 min and the cells harvested by centrifugation at 5000 x g for 10 min at 4 °C. Following this, the pellet was washed once in electroporation buffer (270 mM sucrose, 1 mM $MgCl_2$, 7 mM $NaHPO_4$ [pH 7.4]), resuspended in 5 ml of ice cold electroporation buffer and incubated for a further 10 min on ice. A 0.8 ml aliquot of the cell suspension was then added to 25µl of 0.5 μg μl^{-1} plasmid DNA in EB buffer, transferred to a pre-chilled 4 mm gap electroporation cuvette (BioRad) and held on ice for a further 8 min. A pulse of 1.25 kV (25 μF, 200 Ω) was delivered to the cuvette using a BioRad Gene Pulsar according to the manufacturer's instructions inside the anaerobic cabinet. The cells were held on ice for 10 min before the addition of 8 ml 2 x YTG medium, followed by incubation at 37 °C for 3 h under anaerobic conditions. The culture was then concentrated by centrifugation and resuspended in 100 μl 2 x YTG medium before plating on 2 x YTG agar with the appropriate selection. The plates were then incubated anaerobically at 37 °C for 16 - 96 h.

2.9.9 Electroporation of Plasmid DNA into *C. acetobutylicum* using the Neeson and Titball method

C. acetobutylicum was serially diluted into 5 ml volumes of 2 x YTG medium and grown anaerobically at 37 °C overnight. The lowest two dilutions that showed growth were then used to inoculate 100 ml sterile 2 x YTG medium, which was then incubated anaerobically at 37 °C until an OD_{600} of 0.6 was achieved. The cells were harvested by centrifugation at 6000 rpm for 30 min at 4 ^{o}C (Beckman JA-14). Following this, the pellet was resuspended in electroporation buffer (270 mM sucrose, 1 mM $MgCl_2$, 7 mM $NaHPO_4$ [pH 7.4]) and centrifuged again at 6000 rpm for 15 min at 4 ^{o}C. The cells were then resuspended in 5 ml of ice cold electroporation buffer and incubated for a further 10 min on ice. 1 μg μl^{-1} of plasmid DNA in EB buffer was transferred to a pre-chilled 2 mm gap electroporation cuvette (BioRad) followed by 300 µl of cell suspension. The cuvettes were sealed with parafilm to prevent oxygen from entering the cuvette, put on ice and electroporated outside the cabinet. A pulse of 1.25Kv, 25µF, 100Ω. was delivered to the cuvette using a BioRad Gene Pulsar II according to the manufacturer's instructions. The cells were transferred back into the cabinet and added to 2.7 ml of reduced 2 x YTG followed by incubation at 37 °C for 3 h under anaerobic conditions. The culture was then concentrated by centrifugation and resuspended in 100 μl 2 x YTG medium before plating on 2 x YTG agar with the appropriate selection. The plates were then incubated anaerobically at 37 °C for 16 - 96 h.

2.9.10 Electroporation into *C.acetobutylicum*: using the Mermelstein and Papoutsakis method

This method was used routinely for all knockout strains and is described by Mermelstein *et al.* (1992). *C. acetobutylicum* was serially diluted into 10 ml volumes of CBM medium and grown anaerobically at 37 °C overnight. The lowest dilution that showed growth was then used to inoculate 60 ml of sterile reduced 2 x YTG medium. This was incubated at 37ºC until it reached an OD_{600} of 1.1. 15 ml of the culture was transferred into four 50

ml sterile tubes and covered with parafilm to avoid them becoming aerobic. The cultures were centrifuged at 5000 rpm at 4ºC for 10 mins. The tubes were transferred back into the cabinet and the supernatant was removed. The cell pellets were resuspended with 2.5 ml of electroporation buffer (270 mM sucrose, 5 mM NaH_2PO_4 [pH7.4]). The solution was placed into a single sterile tube, sealed with parafilm and centrifuged at 5000rpm at 4ºC for 10 mins. The buffer was poured off and the pellet was resuspended in 2.3 ml of electroporation buffer. 570 µl of the cell solution was added to a 4 mm electroporation cuvette and 25 µl of 1 $\mu g ml^{-1}$ of DNA in EB buffer was placed into the cell solution and gently mixed (manually without aspiration). The cuvettes were incubated on ice for two minutes. A pulse of 2.0Kv, 25µF, ∞ resistance was delivered to the cuvette using a BioRad Gene Pulsar II according to the manufacturer's instructions. 1 ml of warmed reduced 2xYTG was added to the cell suspension and the cells were incubated for 2 hours. The suspension was then concentrated by centrifugation and resuspended in 800 µl 2 x YTG medium before plating on 2 x YTG agar (200 µl on 4 plates) with the appropriate selection. The plates were then incubated anaerobically at 37 °C for 16 - 96 h.

2.10 Excision of MLS cassette using FLP recombinase

This protocol was used to achieve the removal of the antibiotic marker from pMTL007C-E2 based insertional mutants and the whole *agrA* recombinational mutant. The strain was transformed with pCLF-1 and grown in CBM supplemented with thiamphenicol. The cells were serial diluted and plated onto CBM agar plates supplemented with thiamphenicol. Single colonies were selected and duplicate plated onto CBM plates supplemented with thiamphenicol and then onto CBM plates supplemented with erythromycin. This duplicate plating was done using a cocktail stick and the colony was gently smeared onto the agar to avoid indentation. The plates were left for two days or until growth was seen on the thiamphenicol plates. Colonies which showed little or no growth on the erythromycin plates were selected if their duplicates showed growth on the

thiamphenicol plates. Deletion of the MLS cassette was confirmed by PCR and sequencing.

2.11 T/A Cloning

PCR products were cloned into pCR 2.1 TOPO (Invitrogen) as described in the manufacturer's instructions. This T/A cloning vector has precleaved ends, which have been treated with terminal deoxynucleotidyl transferase to create ddT overhangs on both 3' ends. PCR products can be ligated directly into the vector due to the addition of 3' deoxyadenosine overhangs by the action of *Taq* polymerase during PCR.

2.12 Blue/ White Selection

Blue/ white selection was possible when DNA inserts were cloned into the *E. coli lacZ*α region present on many standard cloning vectors. This leads to disruption of the protein's function, giving rise to a screenable phenotypic marker. Colonies were screened on agar plates containing 40 $\mu g.ml^{-1}$ X-Gal and 1 mM IPTG. Colonies harbouring vector containing the insert grew white, whilst those without grew blue due to a functional β-galactosidase activity and hence uninterrupted gene. This screening process was only possible in *E. coli* strains that carry a mutation in the native *lacZ*α region, namely TOP10, TOP10F' and DH5α.

2.13 RNA Analysis

To minimise RNase contamination, all RNA work was carried out in a designated clean area, which had been thoroughly, cleaned with RnaseZap (Ambion). Certified RNase clean tips and microfuge tubes (Eppendorf) were used and where possible all solutions were treated with 1% (v/v) diethyl pyrocarbonate (DEPC). In addition, all glassware and apparatus were wiped over with RnaseZap and rinsed with DEPC-treated water before use.

2.13.1 Preparation of total RNA from *C. acetobutylicum*

C. acetobutylicum was grown in CBM and samples were taken at an OD_{600} of 0.3 and 0.9. The cells were harvested by centrifugation at 15,700 x g for 2 min in a benchtop microfuge. The supernatant was thoroughly aspirated and the pellet loosened by briefly vortexing. Lysozyme from Ambion's Gram Cracker Reagents kit was diluted in lysozyme dilution buffer (provided by manufacturer) to a concentration of 3 mg ml^{-1}; the pellet was resuspended in 50 μl of this solution by vigorous vortexing and incubated at 37 °C for 10 min. Total RNA was then extracted from the *C. acetobutylicum* lysate using Qiagen's RNeasy Kit, as described in the manufacturer's instructions. Briefly, 350 μl of buffer RLT (supplied with kit) containing β-mercaptoethanol, was mixed with the lysate, followed by the addition of 250 μl of cold absolute ethanol. This solution was then applied to an RNeasy column and centrifuged in a bench top microfuge at 8000 x g for 15 sec. The flow-through was discarded and 700 μl of buffer RW1 (supplied with kit) was applied to the column for 5 min. The column was again centrifuged at 8000 x g for 15 s and the flow-through again discarded. The column was then washed by addition of 500 μl of buffer RPE (supplied with kit) and centrifuged as before. The column was then washed with 500 μl of buffer RPE for a second time and centrifuged at 8000 x g for 2 min. The flow-through was again discarded and the RNA eluted by the addition of 50 μl of DEPC treated water and centrifugation at 8000 x g for 1 min.

2.13.2 Removal of Contaminating Genomic DNA from Total RNA

Contaminating DNA was removed from RNA samples using Ambion's Turbo DNase as described in the manufacturer's instructions. The appropriate volume of 10 x Turbo DNase buffer was added to the RNA sample, 5 U of Turbo DNase per μg of RNA was then added and the final volume adjusted with DEPC-treated water. The reaction was incubated at 37 °C for 30 min.

2.13.3 DNase I Treated RNA Sample Clean Up

Total RNA that had been treated with DNase I was cleaned up prior to use in RT-PCR using Qiagen's RNeasy columns as described in the manufacturer's instructions. Briefly, the RNA sample was adjusted to 100 μl with DEPC-treated water and 350 μl of buffer RLT (supplied with kit) was added to the sample, followed by the addition of 250 μl of cold absolute ethanol. This solution was then applied to an RNeasy column and centrifuged in a bench top microfuge at 8000 x g for 15 sec. The column was then washed by addition of 500 μl of buffer RPE (supplied with kit) and centrifuged as before. The column was washed for a second time by the addition of 500 μl of buffer RPE and was then centrifuged at 8000 x g for 2 min. The flow-through was again discarded and the RNA eluted by the addition of 30 μl of DEPC-treated water and centrifugation at 8000 x g for 1 min.

2.13.4 One-step RT-PCR

Before use in RT-PCR, RNA was treated with Turbo DNase as described in Section 2.13.2 and subsequently cleaned up as described in Section 2.13.3. RT-PCR was carried out using Qiagen's one step RT-PCR kit as described in the manufacturer's instructions. For the *agrD/C* and *agrC/A* transcripts Mg^{2+} solution was added after the 50ºC incubation to make a final concentration of 10 mM. The concentration of the template RNA was obtained using a UV spectrophotometer (A_{260} / A_{280}). The final concentration in the RT-PCR mix was 0.01 $\mu g ml^{-1}$. The reaction was run separately to ensure optimum annealing temperatures and Mg^{2+} concentrations.

RT-PCR products were visualised by agarose electrophoresis with ethidium bromide as described in **Section 2.7.**

2.14 DNA Microarray

2.14.1 Preparation of RNA from *C.acetobutylicum*

The preparation of RNA followed the method described by Tomas *et al.* (2003a). Three cultures of wild type *C.acetobutylicum* and the *agrB* mutant strain were grown in CBM medium to an OD_{600} of 0.6 and 1.4. Cells were harvested by centrifugation at 4°C 5,000 × *g* for 10 min. Pellets were resuspended in 200 µl of SET buffer (25% sucrose, 50 mM Tris-HCl [pH 8], 50 mM EDTA [pH 8]) with 20 mg ml^{-1} of lysozyme, and the samples were incubated at 37°C for 5 min. Cold TRIzol reagent (1 ml; Invitrogen, Carlsbad, Calif.) was added, and the samples were vortexed for 30 s. The TRIzol samples were immediately frozen at −80°C, and the RNA was purified within 1 month to avoid degradation. For isolation and purification, the TRIzol samples were thawed at room temperature and diluted fivefold in ice-cold TRIzol up to 1 ml. Chloroform (200 µl) was added to 1 ml of the diluted TRIzol-treated samples, vortexed, and allowed to stand for 2 min at room temperature. The samples were centrifuged at 12,000 × *g* for 15 min at 4°C, and the aqueous phase was transferred to a fresh tube. Isopropanol (0.5 ml) was added, the tubes were inverted several times, and the samples were allowed to stand for 10 min and then centrifuged at 10,000 × *g* for 10 min at 4°C. The resulting pellet was washed with 75% v/v RNase-free ethanol and spun at 8,000 × *g* for 4 min at 4°C. After drying for 10 min, the RNA was resuspended in RNase-free water and quantitated with a UV spectrophotometer (A_{260} / A_{280}). Each sample was also checked using agarose gel electrophoresis for lack of RNA degradation. Samples were stored at −80°C.

2.14.2 Microarray method

The microarray was performed by Victoria Wright at the Centre for Biomolecular Sciences at the University of Nottingham.

The arrays were printed using the AROS Oligo Set *Clostridium acetobutylicum* V1.0 (Operon) onto UltraGaps slides (Corning Inc) using the MicroGridII arrayer (BioRobotics).

For each array 10µg of RNA was reverse transcribed and directly labelled with Cy5-dCTP and 2µg of gDNA was directly labelled with Cy3-dCTP. Samples were purified and hybridised onto the arrays for 16 h, after which time they were washed and dried. Scanning of the arrays was performed using the Axon 4000B GenePix Scanner, the data extraction software was GenePix Pro 6, both from Molecular Devices (Sunnyvale, USA).

Analysis was performed using GeneSpring GX7.3.1 (Agilent Technologies, Santa Clara, USA). The array data underwent normalisation, where values below 0.01 were set to 0.01. Each measurement was divided by the 50th percentile of all measurements in that sample. Each value was divided by the median of its measurements in all samples. If the median of the raw values was below 10 then each measurement for that value was divided by 10 if the numerator was above 10, otherwise the measurement was discarded.

Conditions were compared to find changes in gene expression according to fold change and Student's T-test p-value cut-off of p=0.01 with Benjamini and Hochberg False Discovery also applied. In the samples taken at an OD_{600} of 0.6, a gene could show higher expression in the *agrB* mutant with respect to the wild type strain at the same point of growth. Alternatively the gene could show higher expression in the wild type strain when compared to the *agrB* mutant at the earlier stage of growth or show no significant difference. These patterns of gene expression were measured in the samples at the later stage of growth (OD_{600} of 1.4). Only genes which showed significant differences in expression (p=0.01) were included in the data.

2.15 Chloramphenicol Acetyl Transferase Assay

C. acetobutylicum was serially diluted into 5 ml volumes of 2 x YTG medium and grown anaerobically at 37 °C overnight. The lowest dilution to grow was then used to inoculate sterile medium (1 in 50 sized inoculum) and allowed to grow anaerobically at 37 °C. A 10 ml sample was removed every hour and the OD_{600} recorded. Each sample was then centrifuged at 5000 x g for 10 min, the supernatant was aspirated and the pellet frozen at -20 °C . The pellets were allowed to thaw on ice, and were resuspended in 1 ml sterile water containing 3 mg ml^{-1} lysozyme by vigorous vortexing, followed by incubation at 37 °C for 30 min with shaking (200 rpm). Following this, the cell suspensions were kept on ice, 200 μl – 250 μl of 150 – 212 μm diameter acid washed glass beads were added and each sample was vigorously vortexed twice for 1 min, with a 30 sec cooling period on ice in between. The lysates were then centrifuged at 10,000 x g for 10 min at 4 °C to remove the cellular debris and beads. The samples were then kept on ice until use. A 100 μl aliquot of lysate was added to 1 ml reaction buffer (10 mM 5,5'-dithio-bis(2-nitrobenzoic acid) in 1 M Tris-HCl [pH 7.8], and 0.1 mM acetyl-CoA) which had previously been equilibrated to 37 °C for at least 30 min, in a clean plastic cuvette. The cuvette was then placed in a spectrophotometer, the A_{412} zeroed and 20 μl of 5 mM chloramphenicol added and quickly mixed by inverting. The reaction was left to run in the spectrophometer until an OD_{412} of 0.1 was reached after which the time it took for the A_{412} to reach 0.2 was recorded. This process was repeated for each sample lysate.

Units per ml of CAT were calculated using the following equation,

$((\Delta A_{412} / \text{min}) / V) \ / 13.6$ where $\Delta A_{412}/\text{min}$ = the increase in A_{412} per min

V = the volume of lysate added

2.16 ClosTron system

Appropriate regions of the gene of interest were selected for targeting and the re-targeting PCR primers were designed using a computer algorithm (Perutka et al.,2004), the use of which is provided as part of the Targetron Gene Knockout System kit (http://www.sigmaaldrich.com).. Appropriate primer IBS, EBS1d and EBS2 sequences for each of the *agr* genes were generated and ordered. The primers for obtaining mutants for *agr*A, *agr*B and *agr*C are shown in Table 2.1. One-tube Splicing by Overlap Extension (SOEing) PCRs were used to assemble and amplify 353 bp PCR products which contain the modified IBS, EBS1d and EBS2 primers. The sequence was verified by cloning into pCR2.1-TOPO (Invitrogen) and using M13 primers and then subcloned as *Hin*dIII / *Bsr*GI fragments into appropriate intron contexts into pMTL007. When using pMTL007C-E2 the PCR product could be ligated directly into this vector and sequenced using 5402F-F1 and 007-R1. The pMTL007 or pMTL007C-E2 based vector was then transformed into *C.acetobutylicum* and transformants were transferred onto CBM containing 2.5µg ml^{-1} erythromycin to select for integrants. Erythromycin-resistant colonies were selected and insertion was confirmed through PCR and sequencing using appropriate primers. The RAM marker was removed from the mutants derived from the pMTL007C-E2 derived mutants using the method described in section 2.10.

2.17 Gas Chromatography Analysis

Ethanol, acetone, butanol, acetic acid and butyric acid were quantified using a Thermo Focus GC equipped with a 30 m TR-FFAP column (0.25 mm internal diameter) and a flame ionization detector (FID). Hydrogen was used as the carrier gas at 0.8 ml/min. The flame was maintained by hydrogen (35 ml/min), compressed air (350 ml/min) and nitrogen (30 ml/min). The injector and detector temperatures were 240°C and 270°C respectively. The column profile used to resolve the peaks was initially 40°C for 2 mins followed by an 80°C/min ramp to 150°C and then a

45°C/min ramp to 210°C and held at 210°C for 1 min. Samples were extracted before injection by adding an equal volume of ethyl acetate (500 μl) to the supernatant sample, vortexing for 10 s and centrifuging for 5 mins at 16,000 x *g*. 300 μl organic phase was removed to a 2 ml sample vial containing a 300 μl deactivated glass insert. 1 μl of sample was injected. A single GC run was used for measuring the samples used. Samples were taken at specific time points from three separate growth curves for each strain used.

2.18 Sporulation assays

2.18.1 Liquid based assay

A single colony was used to inoculate 10 ml of CBM broth. The cultures were allowed to grow for 120 h at 37°C. A 1ml sample was taken out of the work station and centrifuged at 13000 rpm for 1 minute. The supernatant was poured off and the cell pellet was mixed with 250μl of chloroform by vortexing. Each sample was left for 10 minutes and 1 ml of PBS was then mixed with each sample. The samples were then centrifuged at 13,000 rpm for 1 min and as much as possible of the chloroform layer removed. The samples were then centrifuged at 13,000 rpm for 1 min and the PBS was removed and the cell pellet was washed again with 1 ml of PBS. After a final centrifugation at 13,000 rpm for 1 min the PBS was removed and the samples were placed back into the work station. The pellets where mixed with 1 ml of CBM before 100 μl of non-diluted, 1:10 and 1:100 dilutions were plated onto CBM plates . The plates were incubated for 48 h at 37°C under anaerobic conditions. After incubation the number of colony forming units was counted manually with the treated wild type strain as a control.

2.18.2 Plate based assay

C.acetobutylicum strains were incubated on CBM plates for 5 days. Colonies of a similar size were selected and diluted 1:20 with water. A 5 µl sample was then placed on a microslide and 3 random photographs were taken at 400 x magnification and the number of vegetative cells and spores were counted manually. This was repeated twice for each plate resulting in 3 photographs taken for each of the three colonies for each strain. The number of spores was then recorded as a percentage of the total number of vegetative cells.

3.0 IDENTIFICATION AND ANALYSIS OF A PUTATIVE *AGR* SYSTEM IN *C.ACETOBUTYLICUM*

3.1 Introduction

To date there is no established evidence to suggest there is a quorum sensing system in *C.acetobutylicum*. The change from acidogenesis to solventogenesis in late exponential phase has been shown to be associated with pH and acetate concentration but so far signalling via a QS molecule has not been implicated.

There are at present other examples of quorum sensing within the genus *Clostridium*. Several studies have shown that clostridia produce signal molecules. Studies have revealed a two component regulatory system consisting of the *virR* and *virS* genes which has been shown to be involved in the regulation of toxins in *C.perfringens* (Lyristis *et al*., 1994; Shimizu *et al*., 1994). There have been reports of a diffusible signalling substance which appeared to be involved in the regulation of toxin production (Higashi *et al*., 1973; Imagawa *et al*., 1981; Tatsuki *et al*., 1981). This signalling molecule 'Substance A' may have a direct effect on the VirR/VirS system. In mutational studies the activity of substance A appeared in exponentially growing cultures and disappeared in stationary phase (Imagawa & Higashi, 1992).

A *luxS* system has also been reported in *C.perfringens* where it is responsible for the production of AI-2 (Ohtani *et al.,* 2002). A *luxS* mutant showed reduced levels of α-toxin, κ-toxin and θ-toxin. A mutation in *luxS* caused a reduction in the transcription of the θ-toxin gene (*pfoA*) during mid-exponential growth but the genes for α-toxin and κ-toxin were unaffected.

C.difficile late exponential supernatant has been shown to induce bioluminescence when it is used in the *Vibrio harveyi* BB170 *luxS* reporter

strain which responds to AI-2. When the genome of *C.difficile* was analysed a *luxS* homologue was found to be present along with an adjacent regulator/sensor kinase *rolA/rolB* (Carter *et al.*, 2005). A vector (pGC1) containing the *C.difficile luxS* homologue was found to complement the *E.coli* strain DH5α which has a frameshift mutation in $luxS_{Ec}$. When pGC1 was electroporated into this strain it induced bioluminescence in a *V.harveyi* BB170 reporter strain suggesting *C.difficile* does encode an active AI-2 synthase. It was found to be possible to regulate the expression of *rolA* using antisense RNA technology and this resulted in increased levels of AI-2. As a result it is postulated that *rolA* is involved in the negative regulation of AI-2 production. Furthermore a *luxS* homologue has also been found to exist in the *C.difficile* strain CCUG19126 (Lee and Song, 2005). They showed some involvement of LuxS/AI-2 in the regulation of both toxin genes *tcdA* and *tcdB* as well as a holin like gene (*tcdE*).

Early log phase cells were exposed to cell free supernatant for 1 h with BHI as a negative control. Real time PCR was used to look at the expression of these three genes and the amplification curves all shifted left as compared with the BHI control showing earlier expression of the genes. Similar results were seen when early-log *C.difficile* cells were exposed to cell-free supernatant from mid-log *E.coli* DH5*a* carrying pGEM-T-*lux*Scd as compared to control medium from mid-log *E.coli* DH5a carrying pGEM-T . As a result is it suggested that $luxS_{Cd}$/AI-2 may have a role in the pathogenesis of *C.difficile*.

The mid exponential phase cell free culture of *C.botulinum* has also been reported to induce bioluminescence in *V.harveyi* BB170, suggesting this species also makes use of an AI-2 (Zhao *et al.,* 2006).

A degenerated strain of *Clostridium saccharoperbutylacetonicum*, a solvent producing *Clostridium* related to *C.acetobutylicum,* has also been restored to a solvent producing phenotype by a small substance extracted from the WT strain (Kosaka *et al.,* 2007).

With previously described evidence of a variety of potential quorum sensing systems it would not be entirely unexpected to find possible quorum sensing associated genes within the genome of *C.acetobutylicum*. For this reason a search for *agr* orthologues was conducted.

3.2 Results

3.2.1 Exploitation of the *C. acetobutylicum* Genome

The genome of *C.acetobutylicum* ATCC824 was the first *Clostridium* genome to be sequenced and annotated (Nolling *et al.,* 2001). Inspection of the genome reveals 3 orthologues of *agrB*, *agrC* and *agrA*. Further investigation revealed a small ORF (downstream of *agrB*) capable of encoding a polypeptide of 54 amino acids which possessed weak homology to the *S.aureus* AgrD protein (see figure 3.1). This putative encoded peptide contained the invariant cysteine residue which is essential to the formation of the thiolactone ring of an AIP. The sequence also contained a conserved processing-signal motif within the C-terminal region. The complete *agr* region is in the same orientation as that found in *S.aureus* (see figure 3.1) with the genes in *agrB-agrD-agrC-agrA* orientation. This raises the possibility that they may be regulated in a similar fashion as in *S.aureus* with all four genes linked in a polycistronic operon.

Current thinking indicated that these two regions are the only factors needed for processing of the AIP by AgrB in *S.aureus.* In the case of AIP it is possible that an alternative type I signal peptidase, SpsB, is responsible for the removal of the N-terminal amphipathic leader (Kavanaugh *et al*., 2007). This may be the reason for the lack of homology within the N-terminal region between *C.acetobutylicum* and *S.aureus.*

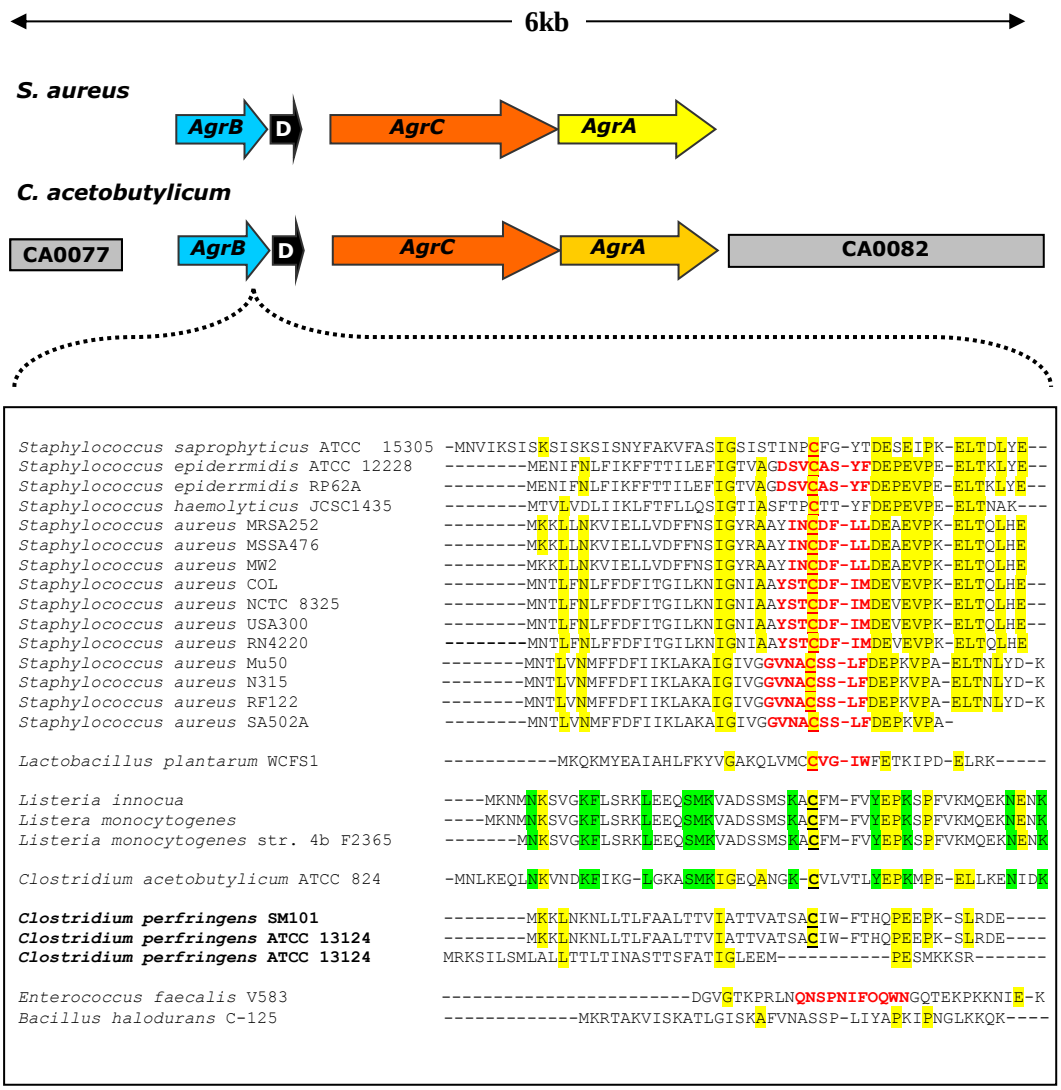


Figure 3.1 The arrangement of the putative *agr* operon in *C.acetobutylicum* and its comparison with the *S.aureus agr* operon. The region encoding *agrD* is highlighted and compared with homologues in other species as previously shown by Wuster and Babu, (2008). The areas in red represent experimentally determined signalling peptides. Yellow highlighted areas show residues which are highly conserved across the homologues and green highlighted areas show areas of homology between *C.acetobutylicum* and *Listeria* species. The underlined C indicates the conserved cysteine residue necessary for the formation of the thiolactone ring.

3.2.2 Transcriptional analysis of the *agr* Operon

Reverse transcription (RT) PCR analysis was used in order to determine what transcriptional linkage was present between all or some of the genes within the operon. RT-PCR is a highly sensitive technique where it is possible to detect a very low copy number of specific RNA molecules. Specific primers are made to anneal to the 3' end of a messenger RNA molecule and the synthesis of complementary (c) DNA is directed in a 3' to 5' direction by reverse transcriptase. This cDNA is then used as a template to synthesise further cDNA strands using conventional PCR.

The experiment used the sensitive OneStep RT-PCR kit which was purchased from Qiagen. Three RT-PCR's were undertaken to produce transcripts 1, 2 and 3 as described in figure 3.2 A/B.

RNA was isolated from samples taken at an OD_{600} of 0.3 and 0.9 representing early and late exponential growth, respectively. The first RT-PCR undertaken (1) was used to assess the transcriptional linkage between *agrB* and *agrD*. As can be seen in figure 3.2 the region between *agrB* and *agrD* is present on a single mRNA species. Furthermore, because the total RNA was calibrated to the same concentration from both the time points the results suggest the genes are transcribed throughout the growth curve. For the RT-PCR 2 and 3 the concentration of total RNA was not calibrated. The concentration measured at OD_{260} for the samples taken at an OD_{600} of 0.3 and 0.9 was 104 ng μl^{-1} and 168 ng μl^{-1}, respectively. This factor may explain the result seen with RT-PCR (2) corresponding to the junction between *agrD* and *agrC* where there is no band seen at the early exponential stage of growth but a band is present at a late stage of exponential growth. These results can possibly be explained by the high sensitivity of the one step RT-PCR kit used and the higher concentration of RNA obtained. Another cause could be aberrant read through from the *agrB* promoter. The most likely reason is probably attributed to an anomalous PCR product. The RT-PCR (3) shows a larger brighter band with the RNA taken from early exponential stage as

compared to the band generated with the RNA from late exponential stage.

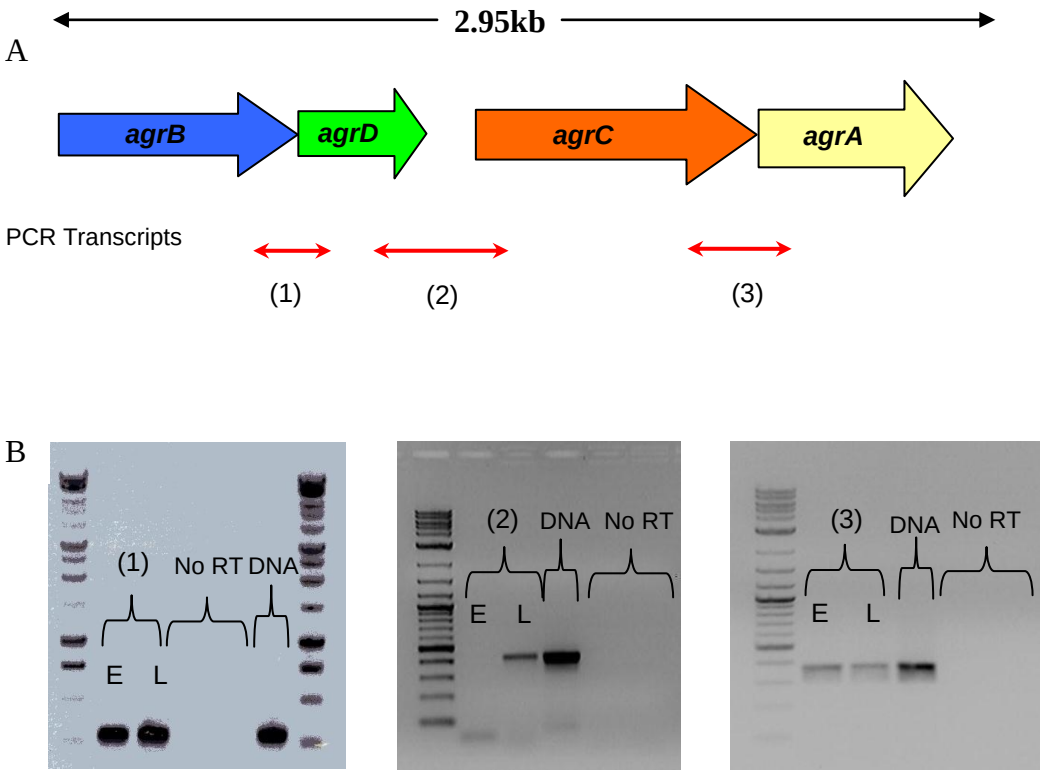


Figure 3.2 RT-PCR analysis of the *agr* locus in *C.acetobutylicum*. *agrB* and *agrD* are predicted to be the two genes responsible for the synthesis of a putative AIP. The downstream orfs *agrC* and *agrA* are thought to encode proteins which form part of a two component signal transduction system. (A) The red bars illustrate the regions amplified by RT-PCR. (B) The RT-PCR was carried out using RNA taken from early-exponential growth (E) and late-exponential growth (L). Control reactions were carried out without reverse transcriptase (No RT) and with a DNA template (DNA). The primers used for junction 1 were agrBDFRT and agrBDRRT. The primers used for junction 2 were agrDCFRT and agrDCRRT. The primers used for junction 3 were agrCAFRT and agrCARRT.

The result suggested that *agrC* and *agrA* may be linked and the genes are expressed at a higher level during early exponential phase. The three RT-PCR's did not confirm that the four *agr* homologues in *C.acetobutylicum* are transcriptionally linked as with the *agr* operon in *S.aureus*. Indeed the results suggested that the two sets of genes (*agrBD* and *agrCA*) are expressed at different levels during growth. For the *agrCA* cluster, the 224

bp region upstream of the start codon of *agrC* is assumed to contain the necessary regulatory elements.

3.2.3 Gene transfer techniques

To date, the transfer of plasmid DNA into *C.acetobutylicum* has only been accomplished using transformation procedures based on electroporation. Successful transfer of DNA is absolutely reliant on the prior methylation of the plasmid due to the presence in *C.acetobutylicum* of a restriction enzyme, CacI824, which cleaves heterologous DNA at the sequence GCNGC. The site can be protected from cleavage using the heterologous methylase Fnu4H1. Practically, this is accomplished by passaging the plasmid DNA to be transformed in a cell carrying a compatible plasmid (pAN1) carrying the gene encoding Fnu4H1 (Mermelstein and Papoutsakis, 1993). Using this system, frequencies of $5.0x10^5$ transformants per ug of DNA have been reported. However, prior to this study, transformation of *C. acetobutylicum* at Nottingham had not been routinely undertaken, and initial experiments failed to reliably achieve transformation. Accordingly, *oriT*-mediated conjugative plasmid transfer from an *E.coli* donor was examined as a possible alternative. This method of plasmid transfer was in routine use at Nottingham for DNA transfer into *C. difficile, C. beijerinckii* and *C. botulinum.*

3.2.3.1 Conjugation of C acetobutylicum 824

To investigate the utility of conjugative plasmid transfer for introducing DNA into *C.acetobutylicum,* an adaptation of the method routinely used for *C. beijerinckii* was explored (see section 2.9.7). For these experiments the plasmid used was pMTL9361, which is based on the replication region on the *C.difficile* plasmid pCD6 and carries a *catP* gene as the selectable marker. In the initial experiments, thiamphenicol resistant colonies were obtained at a low frequency. Equivalent colonies were not obtained when either the donor or recipient alone were plated on selective media.

Subsequent mini-plasmid preparations prepared from a random selection of the low number of chloramphenicol resistant transformants obtained were used in a PCR using primers (CDcond1 and CDcond4) specific to pMTL9361. DNA fragments of the expected size were demonstrated upon subsequent analysis of the reaction on agarose gels (figure 3.3). The additional use of primers directed against the *amyP* gene confirmed the identity of the cells as *C. acetobutylicum* (figure 3.4). The cleared lysates prepared were also used to transform *E.coli* with selection for chloramphenicol resistance.

These experiments appeared to show that conjugation could be used to introduce plasmids into *C.acetobutylicum.* However, two subsequent repetitions of the method failed to yield thiamphenicol resistant colonies. Although it was possible to conjugate a plasmid into *C.acetobutylicum* this was not found to be reproducible. This could have been because of the low efficiency by which plasmids can conjugate into this species as with transformation.

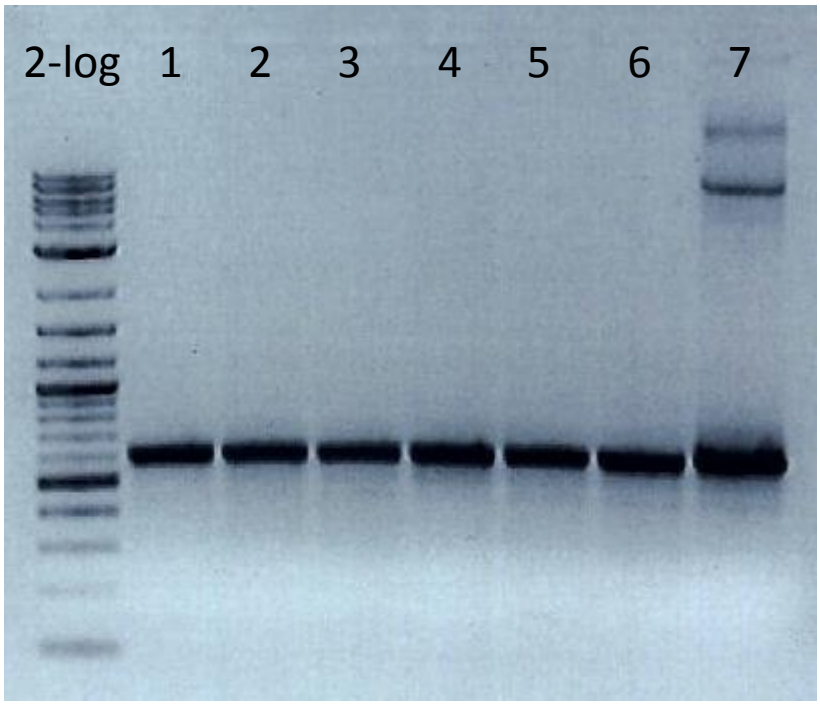


Figure 3.3 PCR products using primers CDcond1 and CDcond4 specific to pMTL9361 and plasmid preparations taken from 6 transconjugants (1 to 6) and plasmid control (7).

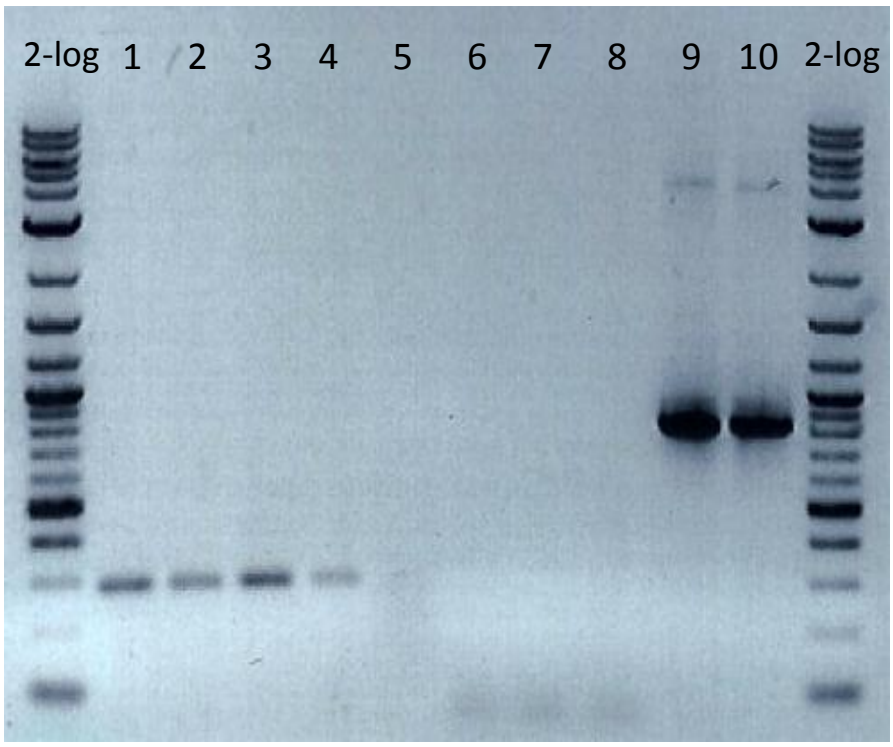


Figure 3.4 An agarose gel showing evidence that 3 of the thiamphenicol resistant colonies which harbour pMTL9361 are *C.acetobutylicum* 824. Genomic preparations were made from 3 of these colonies. Lanes 1 to 3 show PCR products using primers AGRBDF and AGRBDR specific to the *agr*B gene of *C.acetobutylicum* 824. Lanes 4 and 5 are the positive and negative controls using known *C.acetobutylicum* genomic DNA and *C.beijerinckii* genomic DNA and the same primers targeted towards *agrB*. To confirm that the colonies were not *C.beijerinckii*, primers targeted to *C.beijeirinckii gutD* were used to test the transconjugant genomic DNA in lanes 6 to 8. Lanes 9 and 10 are positive controls using the same primers and *C.beijerinkii* genomic DNA.

3.2.3.2 Inhibition of restriction modification system using ARD and its effect on conjugation efficiency

One factor that could be contributing to the variability observed in conjugative transfer of the plasmids is the presence of additional restriction-modification systems over and above CacI824. *C.acetobutylicum* is known to possess at least 6 restriction modification systems. 5 of these are type II systems including MCac824I. The final system, Cac824MrrP, is unclassified. The presence of these systems may contribute to the low frequency of DNA transfer found in *C.acetobutylicum* which can be alleviated by methylation of the DNA before introduction to the host. So far there have been no type I restriction systems identified in *C.acetobutylicum*.

The transfer of DNA through conjugation has been found to be more resistant to restriction than transformation (Read et al., 1992). One of the reasons for this resistance is thought to be due to the presence of an 'alleviation of restriction of DNA' (ARD) system in some conjugative plasmids. The first plasmid antirestriction system to be discovered was the *ardA* system of the IncL1 group of enterobacterial plasmids (Delver et al., 1991; Read et al., 1992). ArdA proteins have been found to inhibit type I restriction systems only. The collB *ardA* functions to alleviate the restriction of the donor plasmid but does not protect plasmids introduced through transformation (Wilkins, 2002). Although a type I restriction modification system has not been identified in *C.acetobutylicum* two plasmids carrying a version of the *ardA* gene were made and introduced into *C.acetobutylicum* 824. These were based on the *ardA* homologues found on the conjugative transposons Tn*916* and Tn*5397* from *B.subtilis* 34A and *Clostridium difficile* 630 respectively. The sequences encoding the *ardA* regions were amplified using PCR. The primer pairs used were TN916F and TN916R and TN5397F and TN5397R.The regions were then cloned into the multiple cloning site within the *lacZ* gene of pMTL9361 to make pJSARD916 and pJSARD5397 as described in figure 3.5. The frequency of transfer of the two resultant plasmids into *C.acetobutylicum* was then tested. No transconjugants were obtained in these *ardA* strains.

This is not an unexpected result as the activity of ArdA in conjugation is thought to be transient due to the brief formation of a novel type of promoter *ssi3* which forms a T-strand during the initial stage of conjugation. This allows the rapid expression of ArdA before the T-strand is converted into duplex DNA (Bates et al., 1999; Wilkins, 2002) The two *ardA* homologues used were under the control of a strong *fac* promoter which may have led to an inappropriate high level of expression. A more probable reason why the *ardA* genes had no effect was due to the fact that *C.acetobutylicum* possesses no type I restriction system including the unclassified Cac824MrrP.

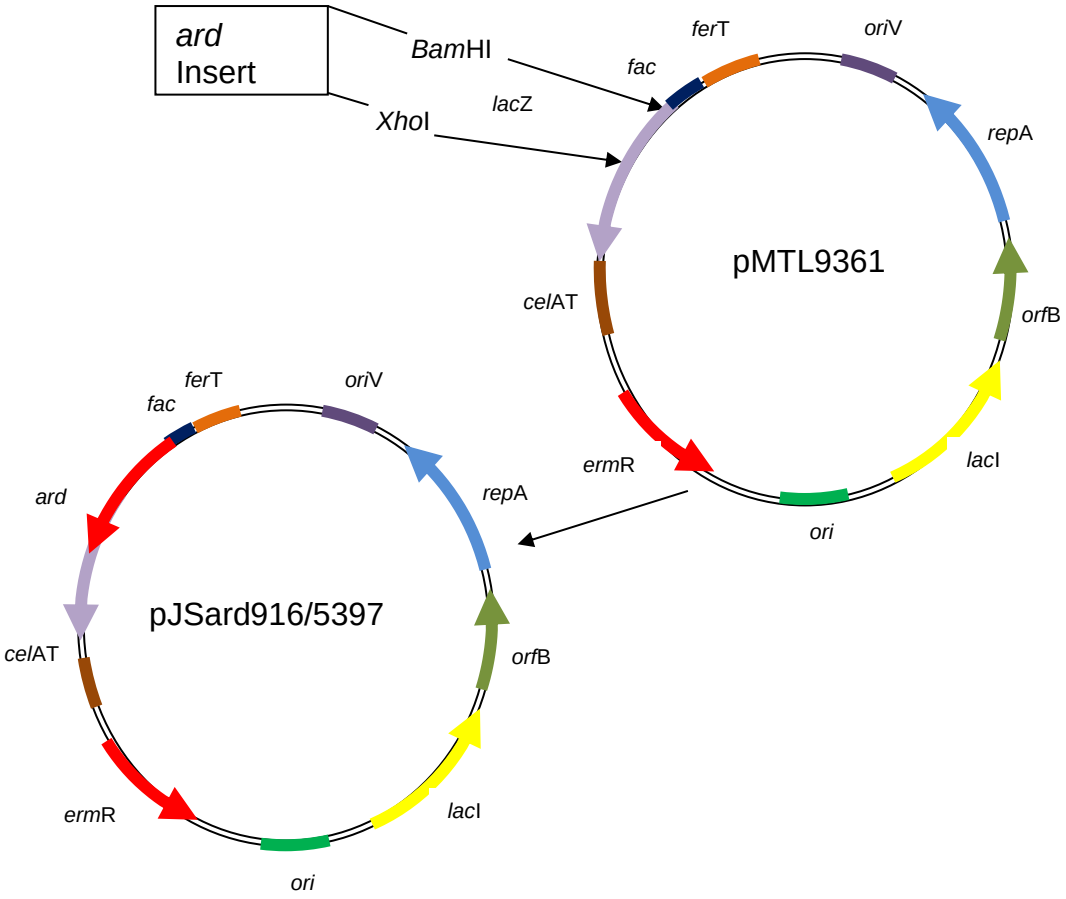


Figure 3.5 Construction of *ard* based plasmids pJSARD916 and pJSARD5397 which were based on the *ard* genes from transposons Tn*916* and Tn*5397*. The PCR was performed using primers TN916F and TN916R for the Tn*916 ard* region and both TN5397F and TN5397R for the Tn*5397 ard* region. Products were checked for the correct sequence and then cloned into the *lacZ* gene of pMTL9361. This has the strong ferredoxin promoter upstream of it and is flanked by two terminator regions.

3.2.3.3 Development of a consistent electroporation protocol

The protocol which was developed and resulted in the most consistent success was based on the protocol used by the Papoutsakis laboratory (Mermelstein et al., 1992). This method described in section 2.9.10 used the same electroporation buffer and parameters as this protocol but used smaller volumes which made the process of preparing electrocompetent cells faster. Although no test was conducted to accurately estimate the efficiency of transformation, the average number of colony forming units was found to be 35 per µg of DNA, sufficient for future studies. Furthermore it was found that cells prepared for electroporation could be stored in 20% glycerol at -80ºC and electroporated once thawed. Further studies are needed to assess if this is a viable method to use routinely.

3.2.4 Construction of reporter vectors

3.2.4.1 GUS based vectors

Knowing that some of the genes within the *agr* locus of *C.acetobutylicum* were transcriptionally linked it was decided to analyse their expression further using reporter gene technology. Two vectors were made for this purpose using pGC4lacIgusAoriT, an inducible expression vector containing the *fac* promoter (figure 3.6). The promoter region upstream of *agrB* was amplified by PCR using AGRBDF and AGRBDR and cloned into the plasmid pCR2.1® TOPO from Invitrogen. The multiple cloning sites which flanked the promoter sequence were then used to digest the plasmid and insert the promoter sequence into pJS01, in effect replacing the *fac* promoter in pGC4lacIgusAoriT. Two methods were initially used for transforming the vector pJS01 into *C.acetobutylicum* 824. The first established by (Mermelstein *et al*. 1992), utilises cells that are taken at late exponential phase while the second method, based on the procedure used for *Clostridium beijerinckii* NCIMB 8052 (Oultram *et al.,* 1988a), uses buffers that are based on magnesium, rather than the sodium phosphate buffer used in the method used by Mermelstein *et al*. (1992). Both methods were used but transformants were not obtained with pJS01 or with the control plasmid, pMTL9361.

It is hard to explain why it was not possible to transform this vector into ATCC824. The plasmid may be unstable in *C.acetobutylicum*. The *gus*A reporter system has been used successfully before in ATCC 824 (Girbal *et al*., 2003) so it is unlikely that the expression of GusA is toxic to the cell. However, it may be possible that the promoter for *agrB* and *agrD* increases the expression of GusA to toxic levels within the cells of *C.acetobutylicum*.

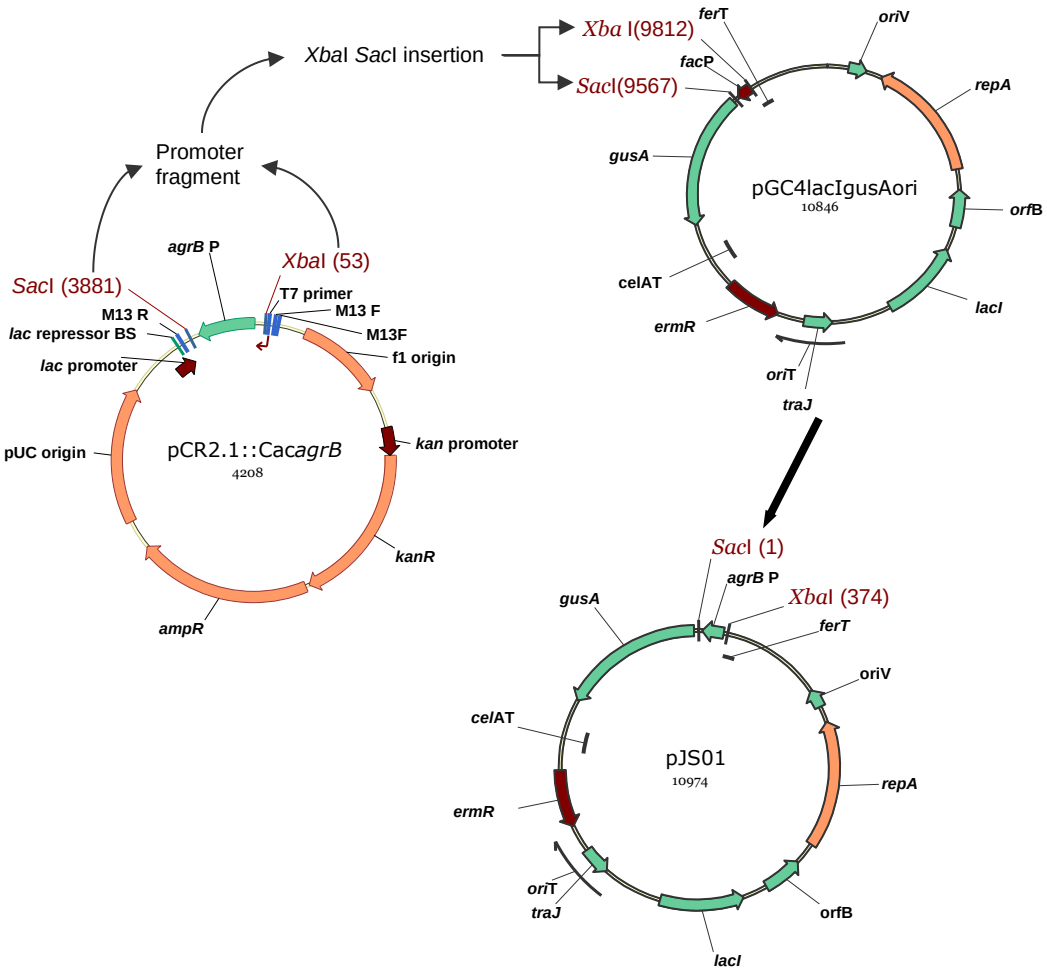


Figure 3.6 Method used for construction of pJS01.

3.2.5 Development of *catP* based expression vectors

Because of problems with the GUS based expression vector an alternative reporter vector was made based on a *catP* gene encoding chloramphenicol acetyltransferase. This system relies on the fact that this enzyme transfers an acetyl group to the antibiotic chloramphenicol which modifies its structure and renders it harmless to the bacterium. This can be exploited using the following reaction:

$$\text{Acetyl CoA + Chloramphenicol} \xrightarrow{\text{(CAT)}} \text{CoA + Chloramphenicol 3-Acetate}$$

$$\text{CoA + DTNB} \rightarrow \text{TNB + CoA derivative}$$

Where:
Acetyl CoA = Acetyl Coenzyme A
CAT = Chloramphenicol Acetyltransferase
CoA = Coenzyme A
DTNB = 5,5'-Dithio-bis(2-Nitrobenzoic Acid)
TNB = 5-Thio-2-Nitrobenzoic Acid

The reaction of DTNB with the –SH group on CoA results in an increase in OD_{412} due to the TNB anion. This increase in OD_{412} over time can be measured making it possible to calculate the enzyme activity and therefore the activity of the promoter.

The putative promoter regions of the *agrB* and *agrC* genes were amplified by PCR using FprimerBDcat/RprimerBDcat and FprimerCAcat / RprimerCAcat primer combinations, respectively. The promoter regions were originally cloned into pCR 2.1 TOPO and then a pGC4 based vector containing the CAT reporter gene *catP* to generate pJS02 and pJS03, respectively (as described in figure 3.7). It was found that these vectors could not be transformed into ATCC 824 using the available methods. As a result the whole fragment containing both the promoter region and *catP* from these vectors was cloned into the *SmaI* site of pIMP1 as described in figure 3.7 to create reporter vectors pJS04 and pJS05. These two vectors and the pIMP1 control were successfully transformed into ATCC 824.

A method for a CAT assay described by Carter, (2006) proved to be unreliable; due to the fact the reaction was not found to be linear over an extended incubation period. Accordingly the CAT assays were modified and performed using the method described in section 2.15. This proved more consistent and easier to conduct than the original referenced method which relies on the increase in the A_{412} over 1 min whereas the modified method described relies on the amount of time it takes for the reaction to proceed from an OD_{412} of 0.1 to an OD_{412} of 0.2. The recorded time was then used to determine the units of activity.

The data obtained (figure 3.8) suggest that the promoter for *agrC* was active in the late exponential stage of growth. This pattern follows the *S.aureus agr* paradigm for the production of *agrC*. A similar pattern was not seen in the case of *agrB*. The assays for both the vectors were repeated three times. As can be seen in figure 3.8 there was a real difference in the number of CAT units per µg of protein although a similar pattern was seen with two of the Three replicates although a much lower level of activity was noted in the third. This could be due to the fact the slight differences in experimental conditions such as temperature and the substrates used.

A microarray study, found *agrB* to be up-regulated in cluster 3 representing the stationary phase of growth (Jones *et al*,. 2008 (additional datafile 2)). One explanation for the lack of agreement with the data from this study may have been that the transformants carrying the two different reporter plasmid constructs could have been inadvertently transposed. Accordingly, plasmid DNA from the two recombinant *C.acetobutylicum* strains was isolated and retransformed into *E.coli*. The plasmids were then subjected to restriction analysis and the inserts where found to have the appropriate restriction pattern and released a DNA promoter fragment of an appropriate size (data not shown).

This suggests that the results obtained with the *agrB* reporter vector either conflict with the current understanding of the expression of *agrB* in

S.aureus or the results obtained here are incorrect due to an error in *cat*P expression. This error could be due to the fact the insert in the vector does not cover the entire promoter region of *agrB* and *agrD* or there was an error in the *catP* gene of the vector. Due to the fact further data from a planned microarray study would confirm or refute these data it was decided not to pursue this further.

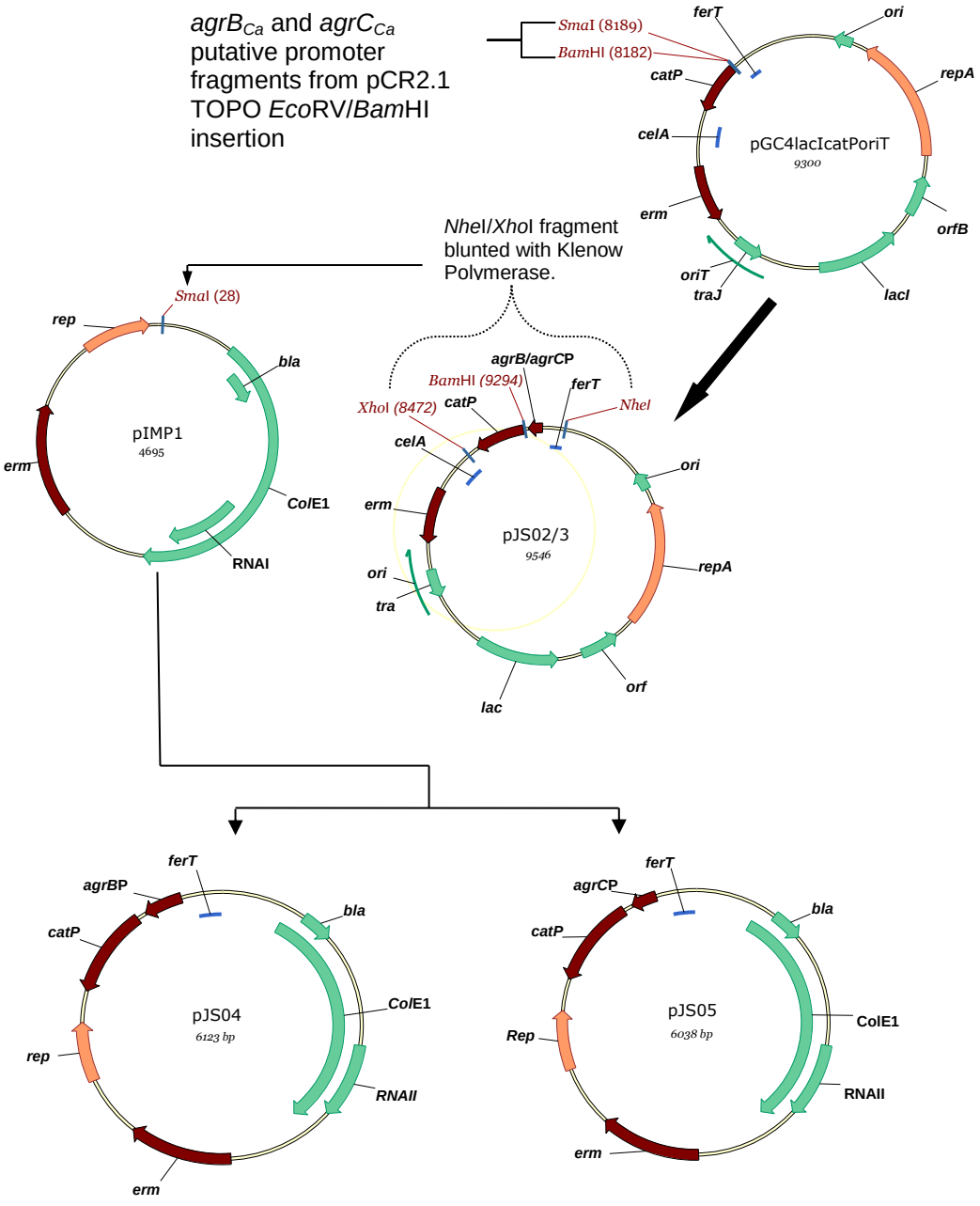


Figure 3.7 Method used for construction of pJS04 and pJS05, the final vectors used for the CAT assay after successful transformation into *C.acetobutylicum*. The original CAT expression vectors pJS02 and pJS03 were unable to be transformed into *C.acetobutylicum*. The promoter and CAT region was therefore inserted into pIMP1, a vector which was compatible with *C.acetobutylicum* to make the final CAT expression vectors.

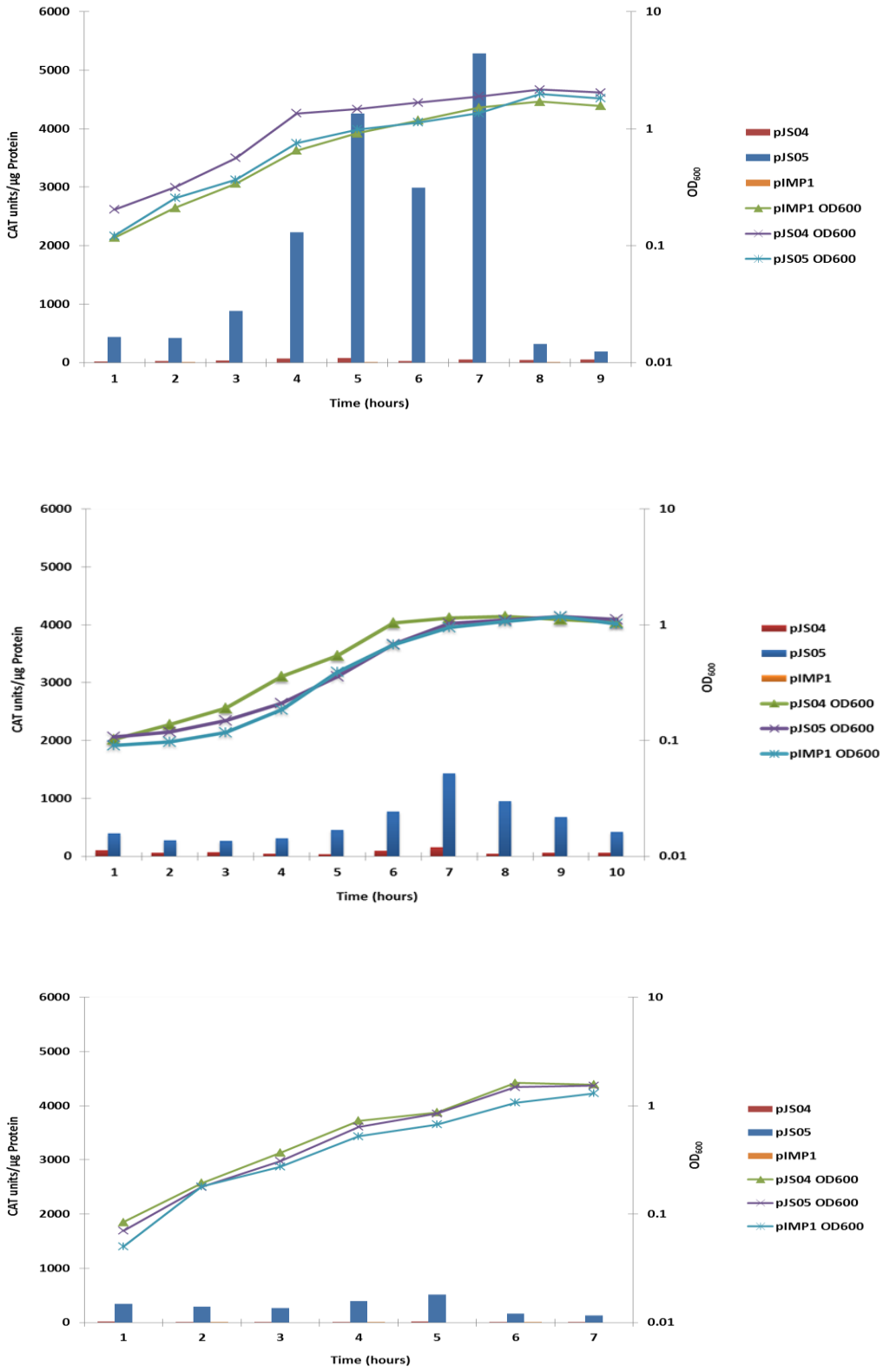


Figure 3.8 Three graphs showing replicate CAT reporter activities for the two reporter vectors. The two vectors used were pJS04 which shows the activity of the promoter region upstream of *agrB* and pJS05 which shows the activity of the promoter region upstream of *agrC*. The CAT activity was so low with pIMP1 it was unquantifiable using this assay.

3.3 Discussion

Previous evidence has suggested that the *agr* system is present in a large number of different firmicutes (Wuster and Babu, 2008). In *S.aureus* the *agr* locus is expressed as a single polycistronic RNA (Kornblum et al., 1990). The evidence presented in section 3.2 suggests the *agr* system in *C.acetobutylicum* produces 2 mRNA species incorporating *agrBD* and *agrCA* respectively The fact that the *agr* system in *C.acetobutylicum* appears to be regulated in a different way does not mean it is not involved with quorum sensing but it may have a different regulatory pathway.

A single PCR product corresponding to the *agrD*C junction was found to be present in the late exponential phase of growth. There was no band present in early exponential growth, however, which suggests the RNA product could be an aberration. It is interesting that the *agr* locus of *C.acetobutylicum* shares close homology with the *agr* locus of *Listeriaceae*. When RT-PCR was performed on the *L.monocytogenes agr* locus the resulting products suggested the co-transcription of the complete *agr* operon and the presence of a full size transcript (Rieu *et al.,* 2007). However, a polycistronic mRNA was never detected by Northern blotting. Further analysis of the *agr* operon using 5´-RACE (rapid amplification of cDNA ends) revealed the presence of 5´ ends which correspond to cleavage and not a response element at the 5´ ends of *agrC* and *agrA*. This suggested post transcriptional processing may have a regulatory role and result in the fine tuning of the individual genes within the *agr* operon of *L.monocytogenes* (Rieu *et al.,* 2007). The fact that there appears to be no response element present in the intergenic region between the *agrD* and *agrC* ORF's of *C.acetobutylicum* suggests that there is no post transcriptional cleavage event in this species. There is also an obvious up-regulation of the promoter activity in late exponential phase of growth in two out of the 3 replicate assays (figure 3.8) as shown with the CAT reporter vector possessing the putative promoter region upstream of *agrC*. This is further proof of an active response element controlling *agrC* and in all probability *agrA*. This also makes it highly improbable that a

polycistronic mRNA containing all 4 ORF's is produced in late exponential phase as suggested by RT-PCR data. A similar expression pattern would be expected with the *agrB* promoter based reporter vector due to the fact *agrB* is involved in the processing of the AIP and therefore should have a similar pattern of regulation. The fact there is no real increase in overall expression can most probably be attributed to an incomplete promoter region.

Various methods of introducing genetic elements into *C.acetobutylicum* were examined and some success was achieved with conjugation. However, this method was found to be a capricious method, proving no more reliable than existing electroporation methods, but it could have some use for further research if it was improved. After the experimental use of a variety of electroporation methods a final reproducible method was developed for general use. Furthermore it was demonstrated that it was possible to freeze the electrocompetent *C.acetobutylicum* cells at -80ºC for future use. This may prove to be a useful benefit in future studies.

4.0 Gene knockout in *C.acetobutylicum* and its effects on solventogenisis and sporulation

4.1 Introduction

In order to study the putative *agr* system in *C.acetobutylicum* further it was necessary to look into disrupting one or more of the *agr* ORF's. The ability to obtain gene knockouts has always been a hindrance when conducting research on the genus *Clostridium*. As a result the understanding of its basic biology has lagged behind that of *Bacillus*. Maximum exploitation of the plethora of determined genome sequences has also been hampered by the lack of available genetic tools for functional genomic studies. Some of the methods described in section 1.17 have been used to try and obtain antisense down regulation or gene knockout in *C.acetobutylicum* ATCC 824.The *agr* system was targeted. Mutants have been obtained in *C.acetobutylicum* ATCC 824 using two different methods. The three genes targeted for mutagenesis were *agrB*, *agrD* and *agrC*. Once these mutants had been obtained the effects of gene knockout on solventogenesis and sporulation were determined using gas chromatography and a spore assay. The methods of analysis had to be developed from existing published methods.

4.2 Results

4.2.1 Development of antisense vectors

At the time of this study the development of knockout systems for *C.acetobutylicum* was in its early stages. The use of antisense RNA strategies for gene inactivation in this species has been widely published. It is for this reason that this strategy was used in the attempt to elucidate the role of *agr* genes in *C.acetobutylicum* by blocking the expression of the *agr* genes.

A great deal of previous work using antisense technology has been conducted by the laboratory of Eleftherios Papoutsakis (Desai et al., 1999; Tummala et al., 2003a; Tummala et al., 2003b, Tummala et al., 2003c). This laboratory used plasmid pSOS95 successfully for antisense RNA work. This plasmid is based on the pIM13 replicon and contains the strong thiolase promoter.

The two regions of both *agrB* and *agrC* encompassing the Ribosome Binding Site (RBS) were amplified using *agrB*asFBam and *agrB*asRBam primers and *agrC*asFBam and *agrC*asRBam primers respectively. These were cloned into pCR2.1-TOPO (Invitrogen) and checked for the required sequence (figure 4.1). The sequence was then ligated into the *Bam*HI site of pSOS95. These vectors were successfully transformed into *C.acetobutylicum* ATCC 824. An RT-PCR was performed with genomic DNA taken from the two antisense strains and they were found to produce the appropriate antisense product.

The development of these vectors was coupled with work to develop a sporulation assay to see if there was any effect of the down-regulation of these genes on sporulation as described in section 4.2.7. A further two vectors were made using the same method described in figure 4.1 but using primers asBD2GENR04/1 and asCA2GENR04/1 as reverse primers with the original forward primers to create pJS08 and pJS09 These two vectors has longer strands of antisense DNA to the target genes. Further work with these antisense strains was stopped and they were not subjected to any biological assays because gene knockout technology became available to target the *agr* genes.

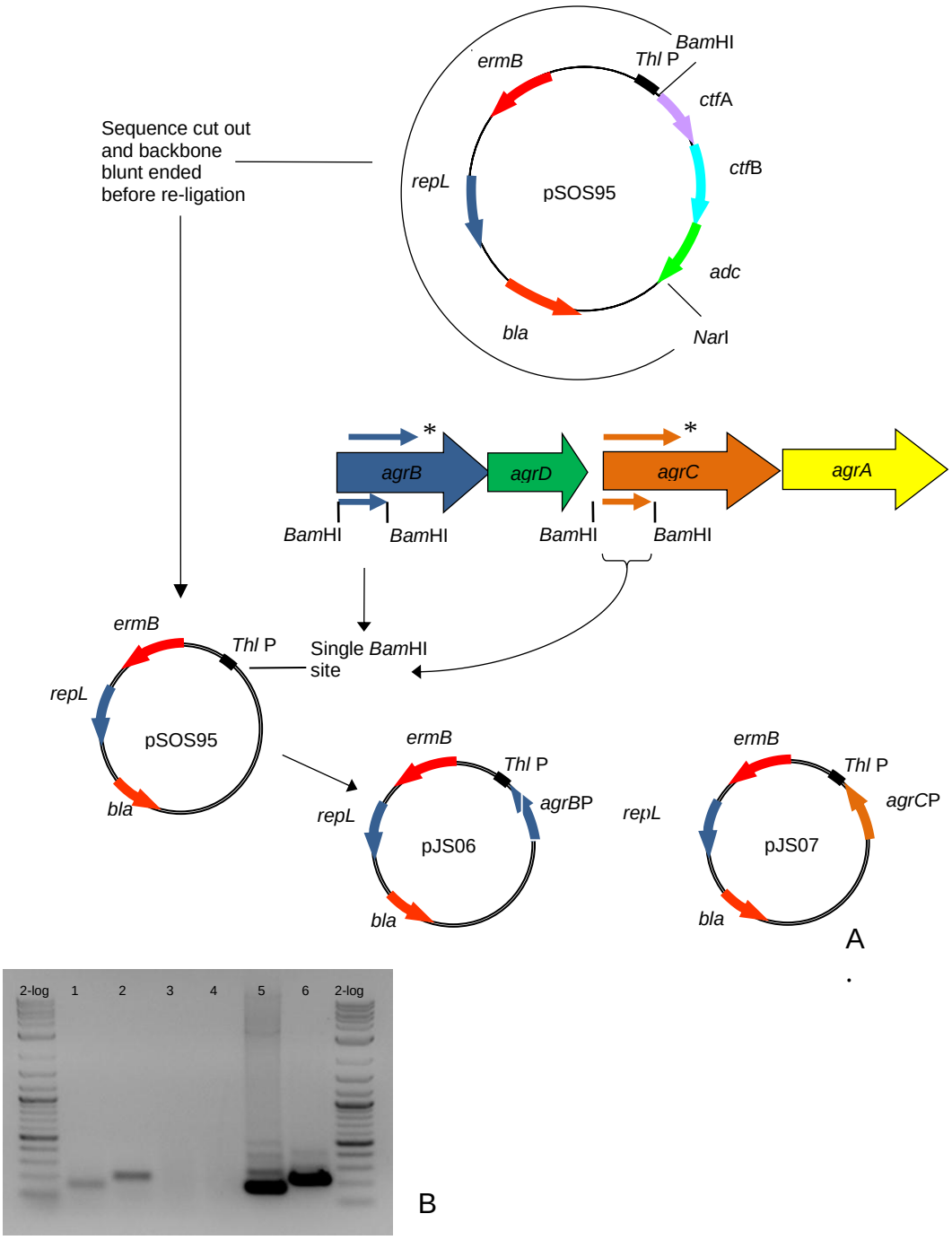


Figure 4.1 Construction of the two antisense vectors pJS06 and pJS07 from pSOS95. The three solventogenesis genes were cut out and the ends were blunt ended to reveal a single *Bam*HI site. The region covering the start of the *agrB* and *agrC* genes which encompass the RBS site were amplified by PCR using primers with a *Bam*HI site. The sequence of the PCR products was confirmed by sequencing using M13 primers after inserting them into pCR2.1-TOPO. They were then inserted into the single *Bam*HI site of pSOS95SS. The two vectors were then selected for the insertion orientated in the antisense direction by restriction analysis. The same method was used for the antisense vectors pJS08 and pJS09 which incorporated the larger PCR products (*). **B**. RT-PCR analysis of pJS06 and pJS07 using forward primers AGRBDF and AGRCAF specific to the region directly upstream from the start codon and reverse primers BDASRTP and CAASRTP specific to a central portion of the antisense region of *agrB* and *agrC*. The RT-PCRs are shown in lanes 1 and 2 and correspond to pJS06 and pJS07 respectively. Control reactions were carried out without reverse transcriptase (lanes 3 and 4) and with a pJS06 and pJS07 template (DNA) (lanes 5 and 6).

4.2.2 Whole gene knockout using allelic exchange

Double crossover integration is a rare event in Clostridia and has only been reported in *C.perfringens* (Awad *et al.,* 1995; Bannam *et al.,* 1995) where a high transformation frequency allows DNA to be efficiently introduced into this organism. This greatly enhances the possibility of detecting such a mutant.

Although antisense RNA has proved to be a useful tool for analysing gene function in clostridia the ideal is to achieve full gene knockout. As previously stated the pursuit of this goal has had limited success. This is particularly true with regards to obtaining a double crossover mutant using allelic exchange.

As previously described, an unpublished method of obtaining double crossover mutants in *C.acetobutylicum* was reported by the group of Phillipe Soucaille at Clostridium IX (Houston). This involved knocking out the restriction modification system (*Cac*824I).

The technique employed (figure 4.2) is typical of standard recombination-based methods for obtaining double crossover integrants. It involves using fusion PCR in order to generate a final product which incorporates extensive regions of DNA flanking the target gene. The target gene itself is replaced with an MLS cassette which itself is flanked by two target sites for the enzyme FLP recombinase. The use of fusion PCR enables the production of the desired sequences with a few PCR reactions.

The Soucaille laboratory kindly donated pCONS2.1 and pREPbuk 3 in order to attempt *agr* knockouts. When the fusion PCR method was put into practice, due to the fact that such large fragments needed to be made the chances of misreading and resulting point mutations was greater. This was found to occur even when using high proof reading Taq polymerase. Since the method required the recombination of large parts of the flanking genes in the organism it was important that these sequences were correct. It was found to be much more efficient to make this resulting fragment

using conventional cloning techniques (see fig 4.2). The resulting plasmid pJS10 was successfully transformed into *C.acetobutylicum* in order to inactivate the *agrA* gene. The method was followed as described in figure 4.2 in order to obtain a double crossover mutant of *agrA*. The deletion was confirmed through PCR. This mutant was used for the analysis of product formation and sporulation in described in sections 4.2.6 and 4.2.7. Despite the success of this method in obtaining a double crossover mutant it was impracticable to repeat this process with the *agrB* gene due to time restrictions.

The next step was to splice the MLS cassette using the pCLF-1 plasmid which contained the FLP recombinase from *Saccharomyces cerevisae*. This vector was transformed successfully into the mutant and the method was followed as described in section 2.10.

The original method described by the Soucaille laboratory was the same as described in Section 2.10 but at the temperature of 37ºC. No evidence for the excision of the MLS cassette was obtained. FLP Recombinase has often been used successfully at a lower temperature with its optimum activity at 30ºC (Buchholz *et al*., 1996). The temperature in the cabinet where the mutants were grown was lowered to this temperature and the method was repeated successfully. The fact that the cells only had to be subcultured a single time indicates that the process is an efficient one. The 50 colonies obtained were duplicate plated on thiamphenicol and erythromycin plates by gently streaking with a tooth pick.

Colonies which grew on the thiamphenicol plate but did not grow with erythromycin or exhibited very little growth were selected. Some of the selected colonies were confirmed to have the erythromycin resistance cassette spliced out by PCR and sequencing using primers agrCasFBam and CONAGRKOR. It must be noted that some of the colonies which seemed to be exhibiting slow growth were also found to have excised the erythromycin resistance cassette which should be noted for future use of FLP recombinase in *C.acetobutylicum*.

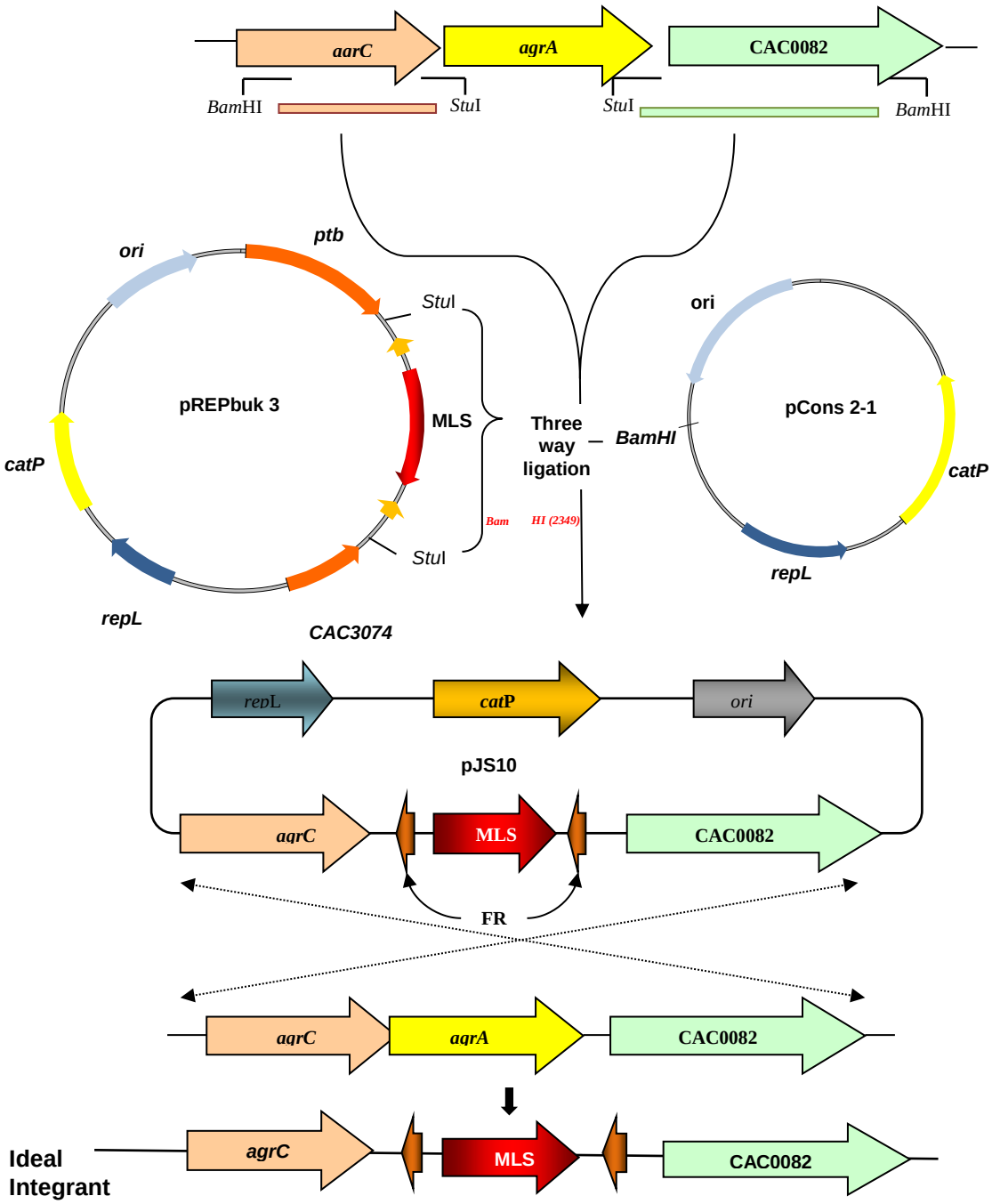


Figure 4.2 Construction of pJS10 used to obtain a double crossover mutant of *agrA* in *C.acetobutylicum*. Two PCR's were obtained of the flanking orfs of *agrA*. The first PCR product was obtained using the agrAKOF1 and agrAKOR1 primers and contained a large amount of the *agrC* sequence. The second PCR product was generated using agrAKOF2 and agrAKOR2 primers and contained the majority of the CAC0082 sequence. The vector pJS10 was constructed using a 3 way ligation of the PCR products which contained the appropriate *Stu*I and *Bam*HI sites, the MLS cassette flanked by FRT sites for FLP recombinase and the pCONS2.1 backbone. The double crossover event was detected by selecting for erythromycin resistant/thiamphenicol sensitive colonies after several subcultures. Deletion of the target gene was confirmed by PCR. The FRT sites are represented by gold arrows flanking the MLS cassette.

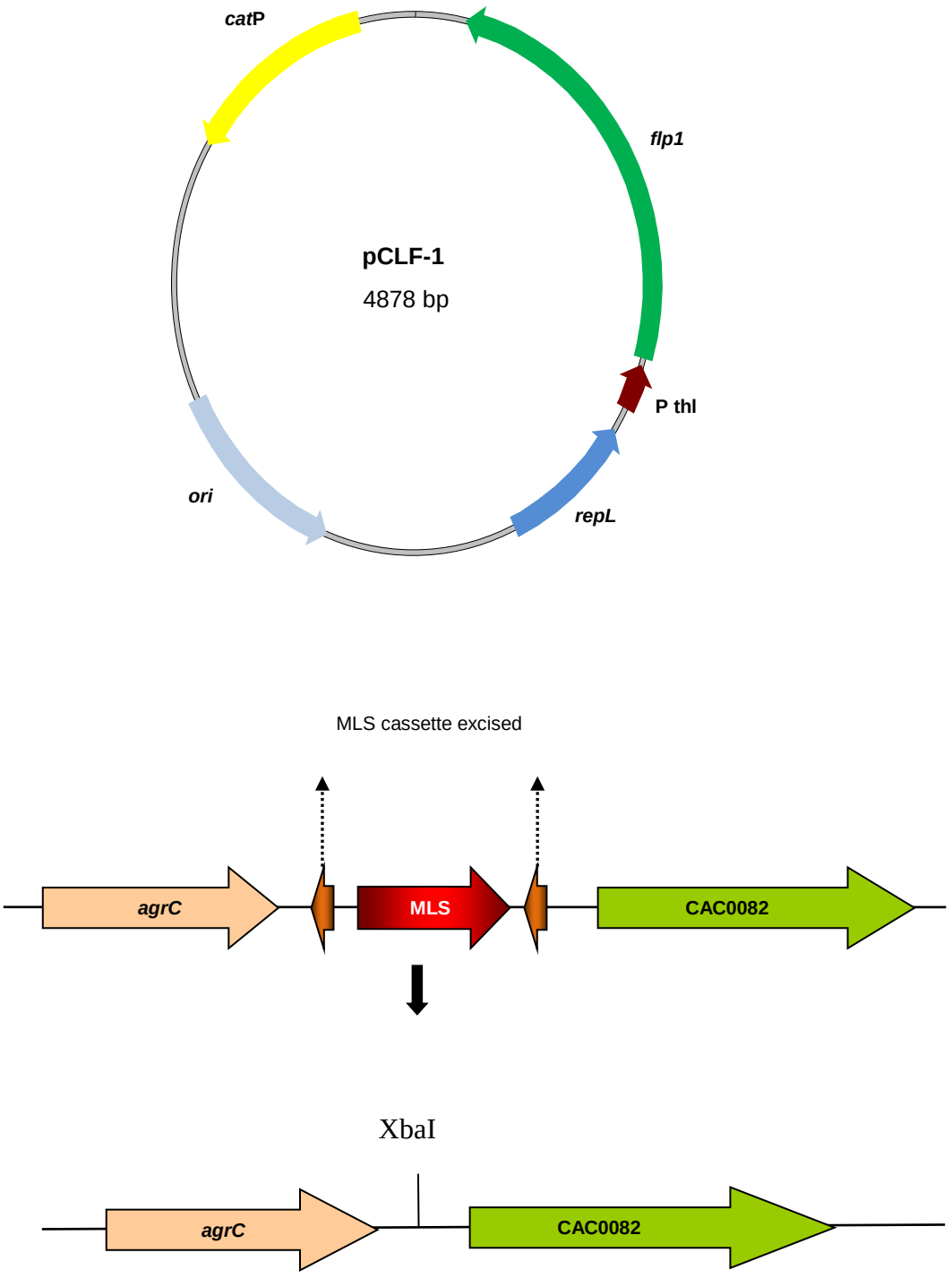


Figure 4.3 The vector pCLF-1 containing the FLP recombinase gene used to excise the MLS cassette from the *agrA* mutant. The excision leaves a single *Xba*I site.

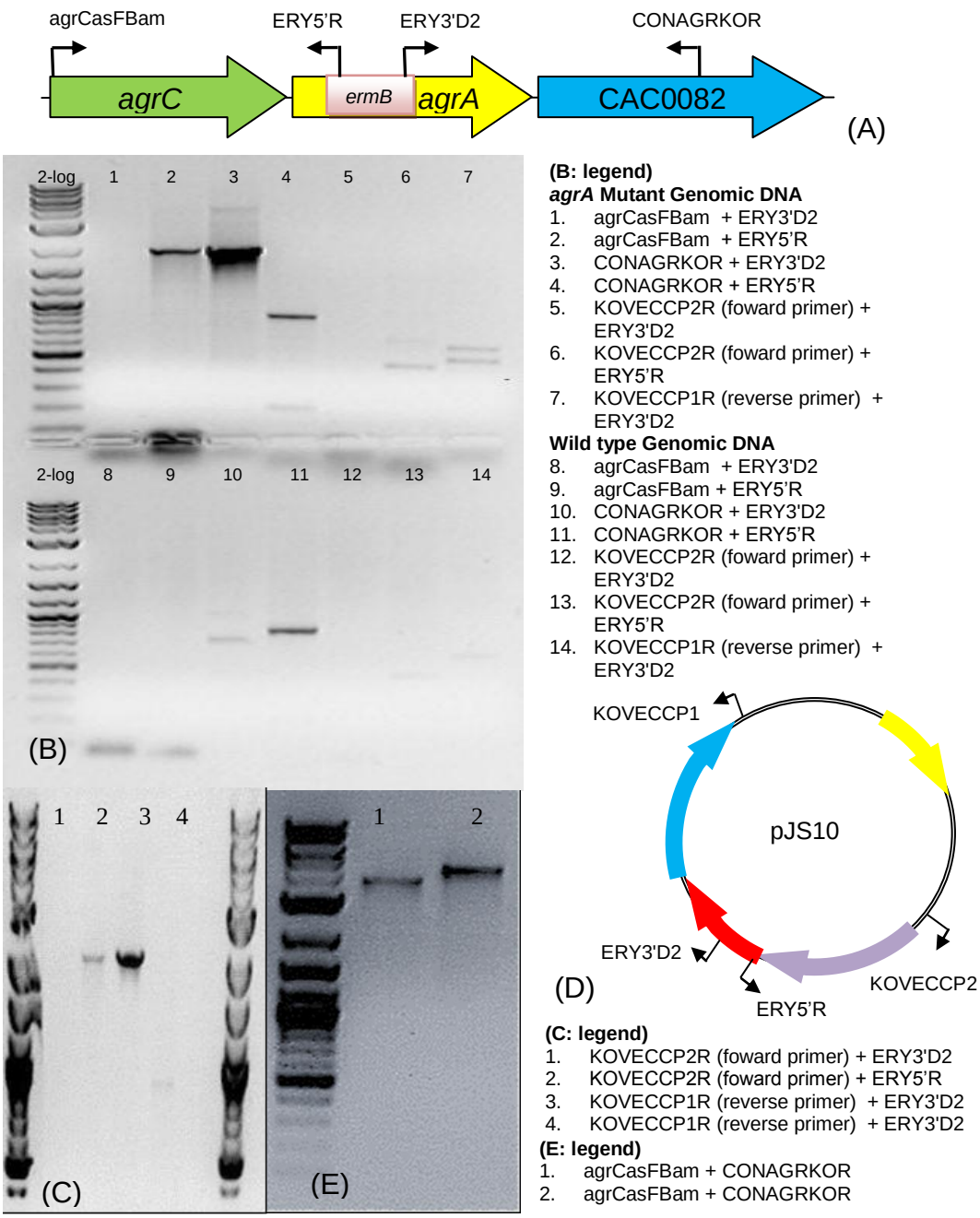


Figure 4.4 Characterisation of the gene inactivation of *agrA* using host mediated double crossover recombination. (A) description of the primer binding sites used for characterisation of the double crossover *agrA* mutant. (B) The first seven lanes show PCR products using a genomic preparation of the mutant DNA. Lanes 2 and 3 show the PCR product using the primers agrCasFBam and CONAGRKOR which flank the region targeted for inactivation. A clear band can be seen in these lanes indicating the presence of the *ermB* selective marker. Lanes 8 to 14 show PCR reactions using wild type *C.acetobutylicum* ATCC 824 genomic DNA. The same bands do not appear in lanes 9 and 10. There is a band in lanes 4 and 11 which could be due to non-specific binding of one of the primers to a second binding site on the genomic DNA. Lanes 6 and 7 show the PCR of the mutant genomic preparation using the primers illustrated in the plasmid diagram (D). 13 and 14 show the same reaction with wild type *C.acetobutylicum* 824 genomic DNA.(C) The positive controls using the primers illustrated in D with pJS10 (lanes 2 and 3) and the mutant genomic preparation (lanes 1 and 4) as a template. (E) PCR products of a reaction using the flanking primers to the site of integration. Lane 1 used the mutant DNA and lane 2 used wild type *C.acetobutylicum* ATCC 824 genomic DNA. The mutant DNA was sequenced using agrCasFBam and CONAGRKOR.

4.2.3 ClosTron Insertional mutation

In order to use gene inactivation to investigate the *agr* system of *C.acetobutylicum* as a whole it was necessary to knock out more than one *agr* gene. The inactivation of the *agrA* homologue using allelic exchange mutagenesis proved to be effective. However, a large amount of time and effort is needed to produce a mutant. This method was not therefore ideal for obtaining the large number of knockouts required. The method developed by Heap *et al*., (2007) was an obvious alternative to use for gene inactivation of the other *C.acetobutylicum agr* homologues.

4.2.3.1 The use of pMTL007 for agr gene knockout in *C.acetobutylicum*

Using the method described in section 2.16 three retargeted pMTL007 based vectors were made designated pMTL007Ca-*agrA*-141a, pMTL007Ca-*agrB*-385a and pMTL007Ca-*agrC*-121s. These vectors were targeted against the *agrA*, *agrB* and *agrC* genes of *C.acetobutylicum* respectively. The three vectors were first methylated by transforming them into *E.coli* Top 10 containing the pAN-2 plasmid created by Steven Cartman (Heap *et al.,* 2007). This plasmid contains the same Φ3T methyltransferase gene as pAN-1 (derived from *B.subtilis* phage Φ3tI) but this was blunt-end ligated into the chloramphenicol resistance marker of pACYC184 which has a tetracycline resistance marker. This enables it to be compatible with pMTL007 based vectors.

The methylated vectors were transformed successfully into *C.acetobutylicum* ATCC 824 and erythromycin resistant colonies were obtained. These colonies were screened by PCR using primers which amplified over an intron-exon junction. Confirmation of the fidelity of the three mutants can be seen in figure 4.5 which shows PCR products using primers flanking the intron-exon junctions. Once a mutant was selected it was then verified by sequencing. The frequency of erythromycin resistant clones obtained which contained the integrated ClosTron was 100%.

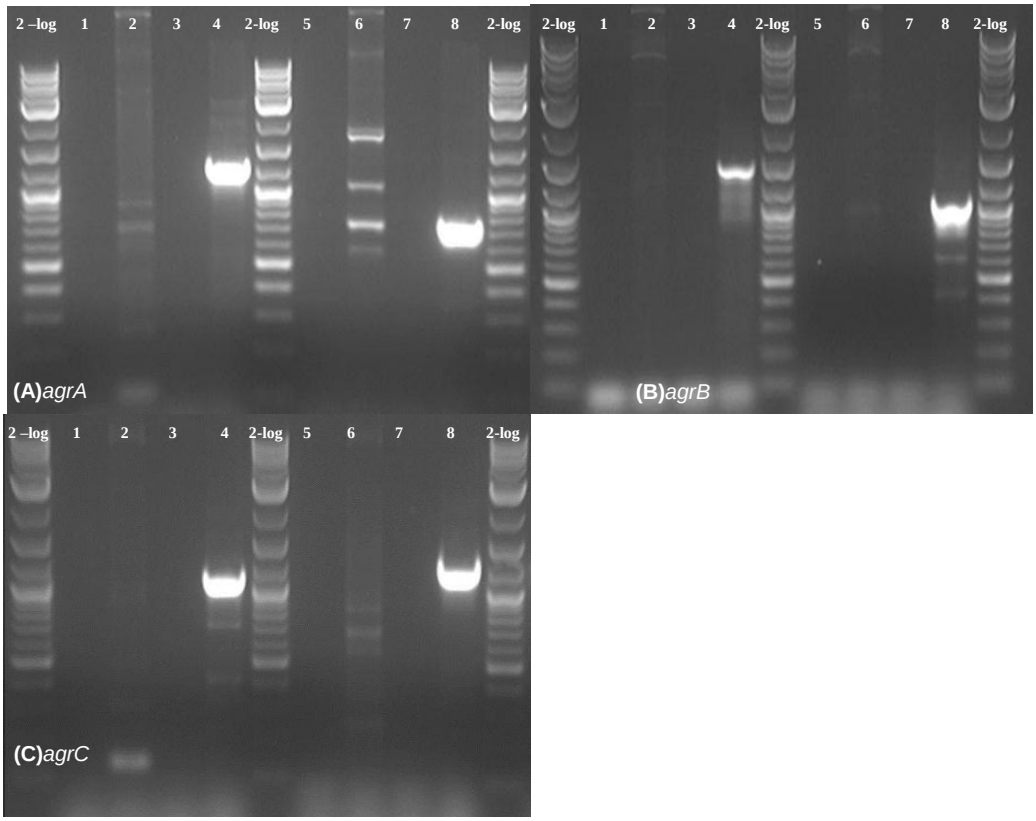


Figure 4.5 Characterisation of the three ClosTron mutants containing the *ermB* RAM using PCR with primers which cross the intron-exon junctions. For all three gel images showing Lanes 1 and 5 are a water control, lanes 2 and 6 are the plasmid control lanes 3 and 7 are *C.acetobutylicum* parental strain genomic DNA and lanes 4 and 8 are the genomic DNA preparations from the respective mutants.(A) The *agrA* insert was inserted in the antisense orientation. For Lanes 1 to 4 the primers used were AGRARVCONF and ErmRam-F. For lanes 5 to 8 the primers used were AgrAasR and EBS universal. (B) The *agrB* insert was inserted in the antisense orientation. For Lanes 1 to 4 the primers used were agrBasF and ErmRam-F. For lanes 5 to 8 the primers used were agrCasF and EBS Universal. (C) The *agr*C insert was inserted in the sense orientation. For Lanes 1 to 4 the primers used were BDASRTP and EBS Universal. For lanes 5 to 8 the primers used were agrCasR and ErmRam-F. The *agrB* mutant was sequenced using SEQAGRBPF and SEQAGRBPR. The *agrC* mutant was sequenced using SEQAGRCPF and SEQAGRCPR. The *agrA* mutant was sequenced using SEQAGRAPF and SEQAGRAPR. All three figures show a PCR product corresponding to the region bodering the insert and the parental DNA. This product is not found in the controls and the wild type strain indicating an insertional mutation.

4.2.3.2 Complementation of mutants

In order to complement the mutants the region encompassing both *agrB* and *agrD was* amplified by PCR using the primers AGRBDF and BamXhoPR1express. This fragment was cloned into pCR2.1 TOPO (Invitrogen) and its sequence confirmed to be correct.

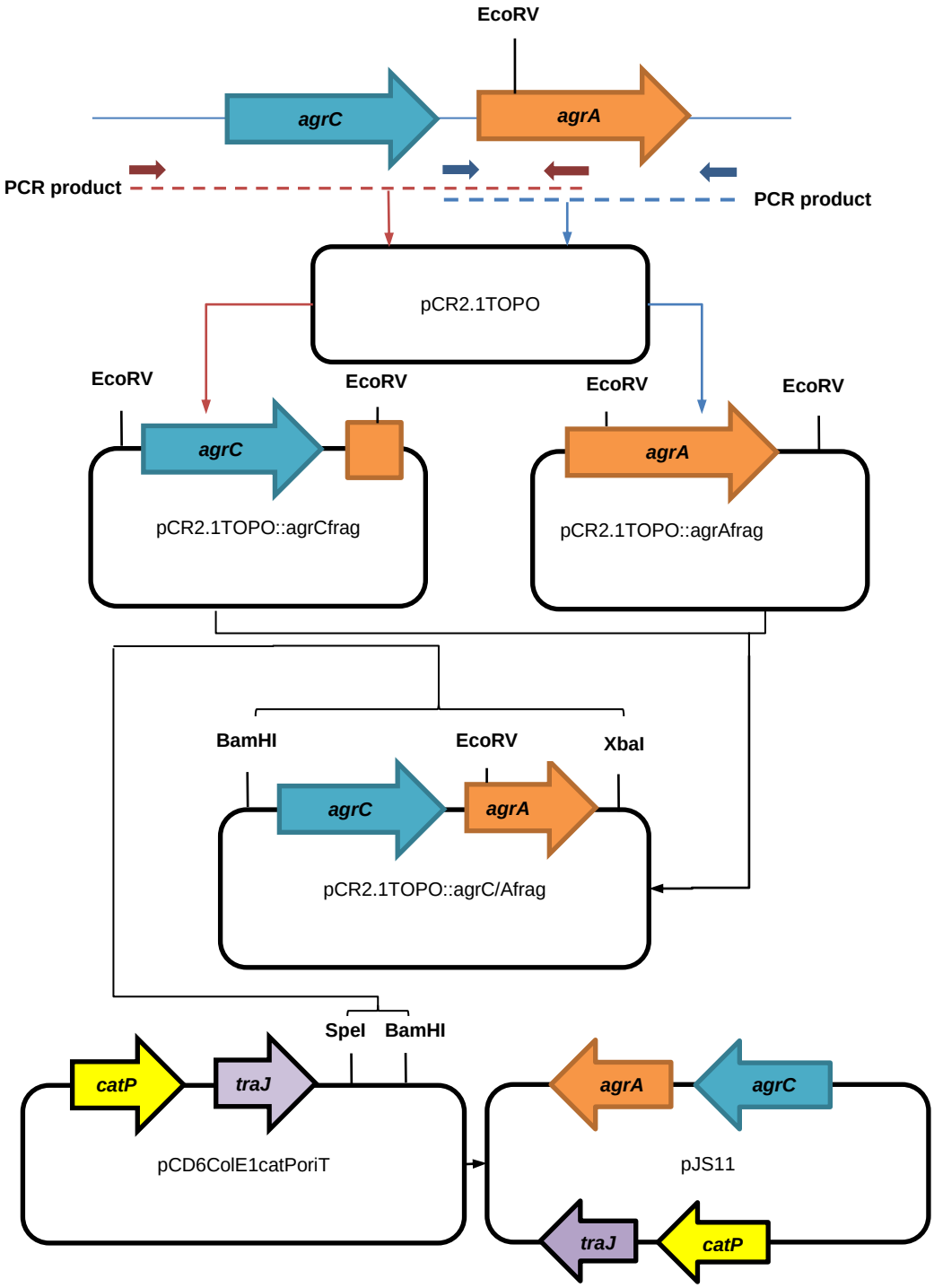


Figure 4.6 Construction of pJS11 which involved a three way ligation of the *agrC* and *agrA* fragments with the pCR2.1 TOPO background to obtain pCR2.1TOPO::agrC/A frag. The whole sequence including the *agrC* and *agrA* genes were sequenced with M13 primers again to ensure the sequence was correct before insertion of the whole sequence into pCD6ColE1catPoriT to get pJS11.

The fragment was digested with *Bam*HI and *Xba*I and ligated into the shuttle vector pCD6ColE1catPoriT using the *Bam*HI and *Sma*I sites to make pJS12. The region encompassing *agrC* was amplified using primers agrCAF and agrARVCONF and the *agrA* gene was amplified using the

primers agrARVCONF and RVNAR1agrCARP. The fragments were cloned into pCR2.1-TOPO (Invitrogen) and the sequence confirmed to be correct through sequencing with M13 primers. The complete method for construction of the pJS11 complementation vector containing the *agrC* and *agrA* regions is described in figure 4.6.

4.2.4 The effect of gene knockout with RAM on *C.acetobutylicum* product formation

The first step of obtaining three *agr* mutants was successfully repeated to show the method was reproducible in *C.acetobutylicum*. The next step was to investigate the mutant phenotypes. With the lack of gene knockouts obtained in *C.acetobutylicum* and the genus as a whole it is understandable that there are few phenotypic assays available without the need for complex equipment. However, there has been a large amount of research into solventogenesis in *C.acetobutylicum* for the purpose of metabolic engineering.

The formation of solvents in *C.acetobutylicum* is one of the most extensively studied aspects of this organism's physiology. The formation of solvents is discussed chapter 1 section 1.4 and there have been many genes shown to affect product formation.

In most laboratories the analysis of organic solvents is undertaken using gas chromatography. As a result this method of analysing the products of the *C.acetobutylicum agr* mutants was used. The method was largely developed by Claire Cooksley for the analysis of ethanol, butanol, acetone, butyrate and acetate and is described in section 2.17. The analysis was undertaken with the wild type *C.acetobutylicum*, and all three of the *agr* mutants and complemented strains. The growth curve for each strain was repeated three times as shown in figure 4.7. All three strains showed similar patterns of growth during exponential and stationary phase although there was some difference in the late stationary phase of growth, particularly with the complemented *agrB* mutant. The average levels of product formation are shown in each panel as a percentage of the average

wild type levels of production at the same time point. The first three time points are shown on the charts along with the 72 hour time point. These are the stages where the products of acidogenesis and solventogenesis would be expected to be produced. The standard error of the mean was used to determine the error bars. A Mann-Whitney *U*-test was used to determine the *z*-score of each mutant and control strain when compared to the wild type strain. This non-parametric test was used because of the small sample sizes. It involves placing all the data in numerical order and then calculating how many data points are not in the hypothesised order. With this test the result is significant only if the absolute value of the z-score was greater than the absolute value of *z* critical. If a 95% degree of certainty is required then $\alpha = 0.05$. For a two-tailed test, the NORMSIN(a/2) function in Excel determines the *z* critical (1.960). The results are significantly different if the z-score is further from the mean than z critical. A *z*-score of less than 1.960 would indicate there was no significant difference between the wild type and the mutants, complemented strains or plasmid controls. The *z*-scores are tabulated in appendices I and II.

4.2.4.1 Effect of *agr* knockout on ethanol production

Figure 4.7 shows the levels of ethanol produced by the three mutants and control strains as a percentage of wild type production. Ethanol should be higher in the later stages of growth after the initiation of solventogenesis (Jones and Woods, 1986). The mean amount of ethanol produced by the wild type went from 6 mM to 10 mM in the first 6 h of growth. The levels peaked at 13 mM after 12 h and decreased again to 10 mM at 72 h growth. The *agrA* mutant shows higher levels than the wild type strain at 2 h which decreased to 20% lower levels than wild type after 6 h ($z = 1.964$). The levels for all three strains were similar to wild type levels at 72 h (*z*-scores < 1.960). There was little difference between the *agrA* mutant and the wild type, complement strain and plasmid control strain, at each time point.

The *agrB* mutant, and plasmid control strain showed a lower level of ethanol from 2 to 6 h when compared to the wild type levels. The complemented mutant showed higher levels of ethanol than the wild type strain at the early stage of growth. At 72 h the *agrB* mutant and the control strains showed higher levels of ethanol compared with the wild type strain although the standard error of the mean was found to be large (z = 1.964).

The *agrC* mutant, the complemented strain and the plasmid control strain all showed lower levels of ethanol than the wild type strain at the earlier stages of growth. Only the plasmid control strain showed higher levels of ethanol than the wild type strain at 72 h. All three strains had a z-score of 1.964 except the *agrC* strain at 6 h.

4.2.4.2 Effect of *agr* knockout on acetone production

Figure 4.8 shows the levels of acetone produced by the three mutants and control strains as a percentage of wild type production. Acetone should be higher in the later stages of growth after the initiation of solventogenesis (Jones and Woods, 1986). In the first six h of growth the mean amount of acetone produced was 3 mM which increased to 8 mM at 72 h. There were higher levels of acetone observed with the *agrA* mutant, its complementation strain and the plasmid control strain at 2 to 6 h of growth when compared to wild type. At 72 h of growth the *agrA* mutant and the complemented strain showed slightly lower levels acetone production when compared to the wild type strain. The plasmid control strain showed similar levels to the wild type strain. At 72 h all three strains had a z-score of 1.964.

A similar pattern was seen with the *agrB* mutant along with the complemented strain and plasmid control strain at the same stages of growth.

The *agrC* mutant and the complemented mutant showed much higher levels of acetone than the wild type strain at the earlier and later stages of growth although the plasmid control strain showed lower levels at the 6 and 72 h stages of growth.

4.2.4.3 Effect of *agr* knockout on acetate production

Figure 4.9 shows the levels of acetate produced by the three mutants and control strains as a percentage of wild type production. Acetate should be at higher levels during the early exponential stages of growth during the onset of acidogenesis (Jones and Woods, 1986). The wild type strain showed produced 11 mM of acetate at 2h, 18mM at 4 h and 31 mM at 6h. At 72 h the concentration of acetate present was 38mM with the highest concentration of 56 mM produced after 24h growth. The *agrA* mutant showed similar levels of acetate to the wild type strain at the earlier stages of growth which is reflected in the low *z*-scores produced. The *agrA* complemented mutant and plasmid control strain showed declining levels of acetate compared with the wild type strain over the earlier stage of growth. The three *agrA* strains all showed higher levels of acetate than the wild type strain at 72 h of growth.

The *agrB* mutant showed lower levels of acetate when compared to the wild type strain at the early stage of growth but levels were similar to the wild type strain at 72 h of growth. No acetate was detected at the early stages of growth with the *agrB* plasmid control strain but levels were similar to the wild type strain at 72 h. The *agrB* complemented mutant showed higher levels than the wild type strain at 4h which may not be significant as shown by the standard error and the *z*-score. The levels of acetate produced by the same strain were similar to the wild type strain at the later stage of growth.

No acetate was detected at the earlier stages of growth with the *agrC* mutant and plasmid control and the only acetate detected with the complemented mutant was at 6 h. The levels at this stage were lower than

the wild type strain. The three *agrC* strains showed similar levels of acetate to the wild type strain at 72 h with all three *z*-scores less than 1.960.

4.2.4.4 Effect of *agr* knockout on butyrate production

Figure 4.10 shows the levels of butyrate produced by the three mutants and control strains as a percentage of wild type production. Butyrate should be at higher levels during the early exponential stages of growth during the onset of acidogenisis (Jones and Woods, 1986). The wild type strain produced 11mM of butyrate at 2h, 24mM at 4h and 43mM at 6h. In the later stages of growth the wild type strain had the highest concentration of 86 mM at 24h which decreased to 70 mM at 72h. The *agrA* mutant, the complemented mutant and the plasmid control strain all produced less butyrate than the wild type strain at the earlier stages of growth. The three *agrA* strains showed similar levels to the wild type strain at 72 h.

The *agrB* mutants showed a similar pattern to the *agrA* strains except the *agrB* complemented mutant which showed higher levels of butyrate than the wild type strain at the earlier stages of growth.

The *agrC* strains showed a similar pattern to the *agrA* strains at the same stages of growth.

4.2.4.5 Effect of *agr* knockout on butanol production

Figure 4.11 shows levels of butanol produced by the three mutants and control strains as a percentage of wild type production. Levels of butanol should be higher at the later stage of growth after the initiation of solventogenesis (Jones and Woods, 1986).The wild type strain produced 3 mM of butanol at 2 h and 4 mM at both 4h and 6 h. At the later stages of

growth the maximum 15mM was produced at 24 h and 13 mM was produced at 72h.

The *agrA* mutant showed lower levels than the wild type strain at the early stages of growth. There was no butanol detected with the *agrA* complemented mutant at the early stage of growth. The *agrA* plasmid control strain showed similar levels of butanol to the wild type strain at the early stage of growth. All three *agrA* strains had similar levels of butanol to the wild type strain at 72 h.

The *agrB* mutant had lower levels of butanol at 2 h of growth than the wild type strain but levels were equivalent to the wild type strain by 6 h of growth. The *agrB* complemented strain had higher levels of butanol than the wild type strain and the plasmid control strain had similar levels to the wild type strain at the earlier stages of growth. All three *agrB* strains had higher levels of butanol than the wild type strain at 72 h of growth.

The *agrC* mutant had lower levels of butanol than the wild type at the earlier stages of growth but the *agrC* complemented mutant and plasmid control strain had higher levels than the wild type strain at 6 h of growth. The *agrC* mutant and plasmid control strain had similar levels of butanol to the wild type strain at 72 h of growth. The *agrC* complemented mutant showed 20% less butanol than the wild type strain at 72 h of growth. All three of the *agrC* strains, the *z*-scores were less than 1.960 at 72h.

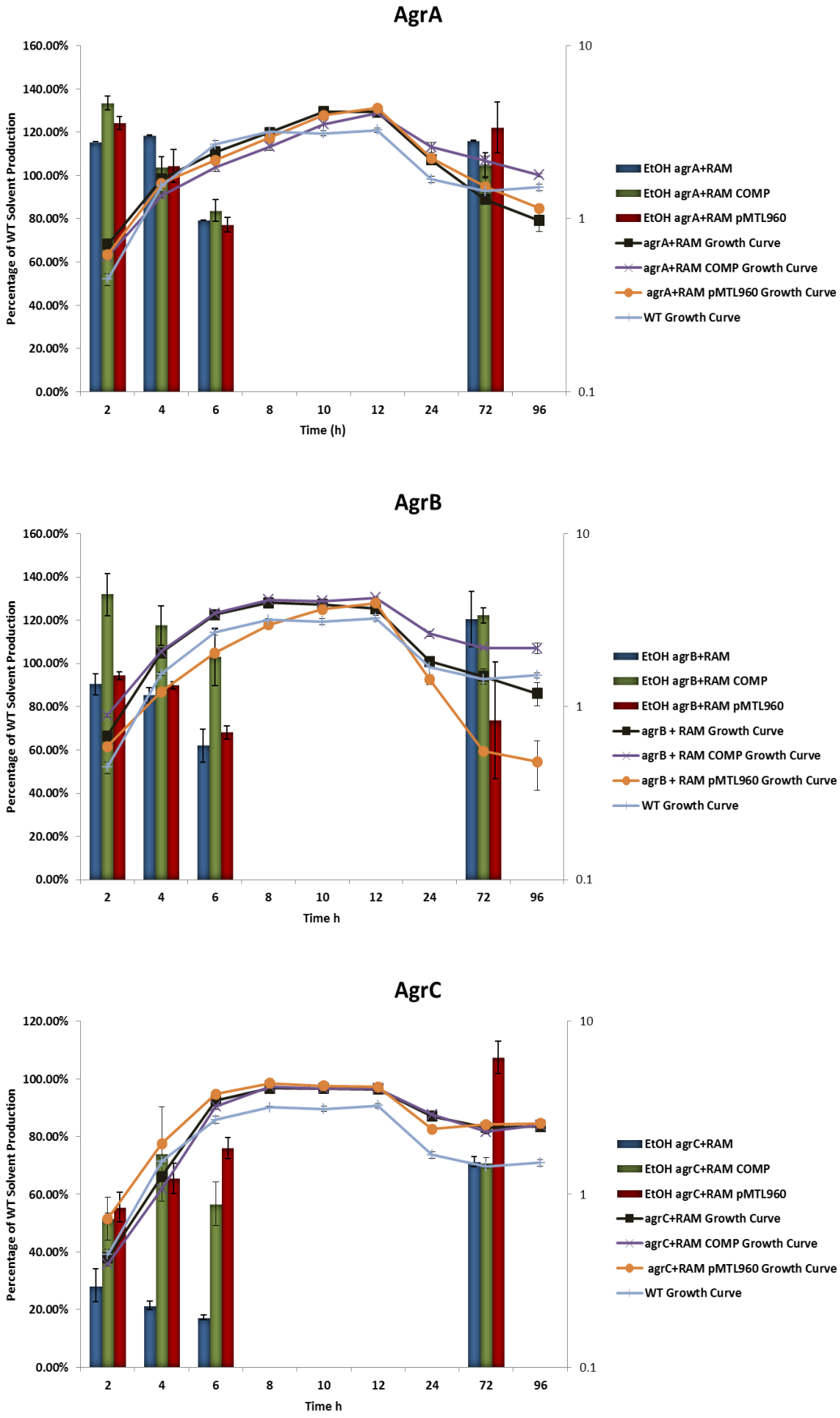


Figure 4.7 Graphs showing the ethanol production of the three *agr* mutants with the RAM inserted as a percentage of WT production at the same stage of growth

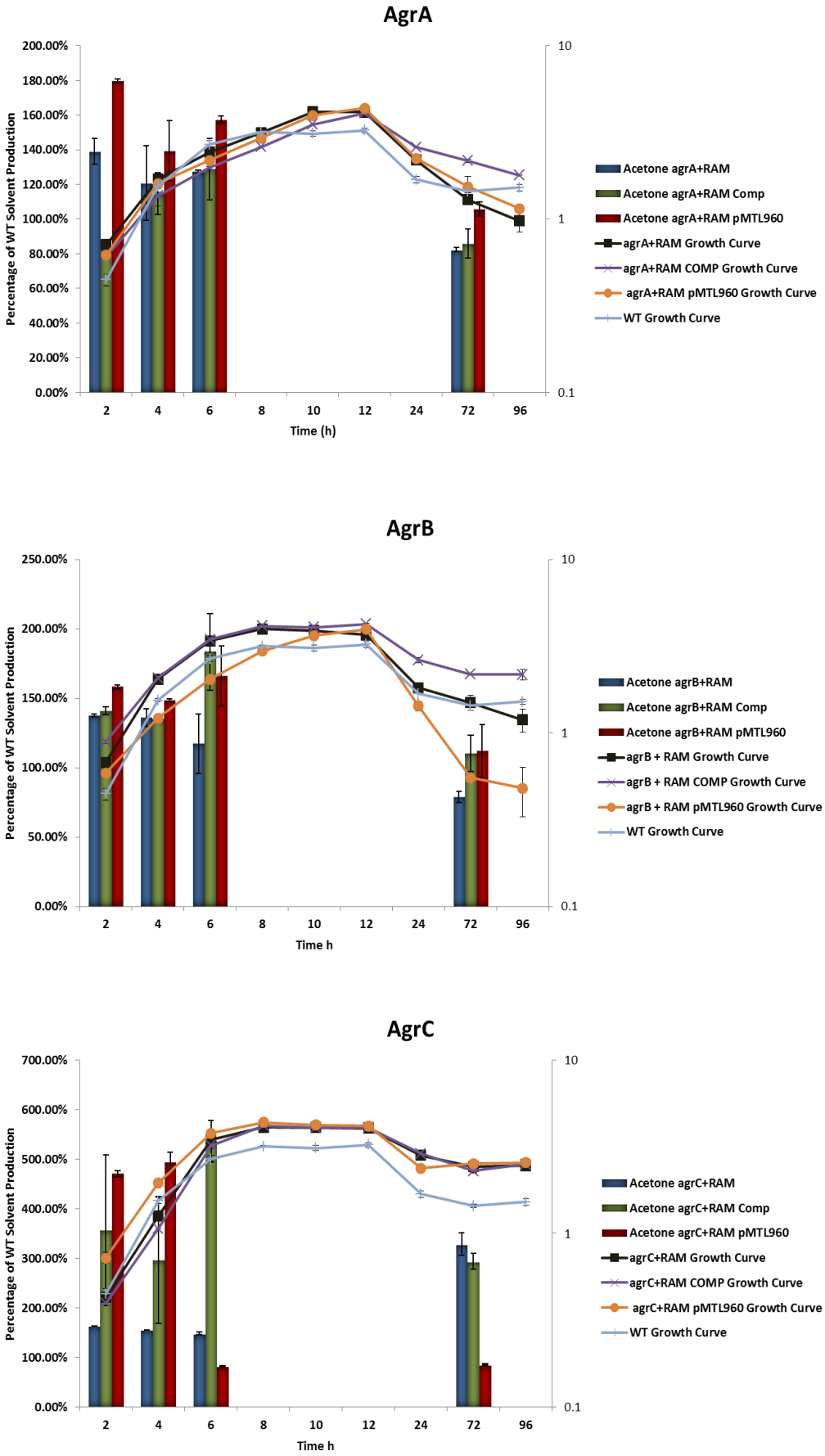


Figure 4.8 Graphs showing the acetone production of the three *agr* mutants with the RAM inserted as a percentage of WT production at the same stage of growth.

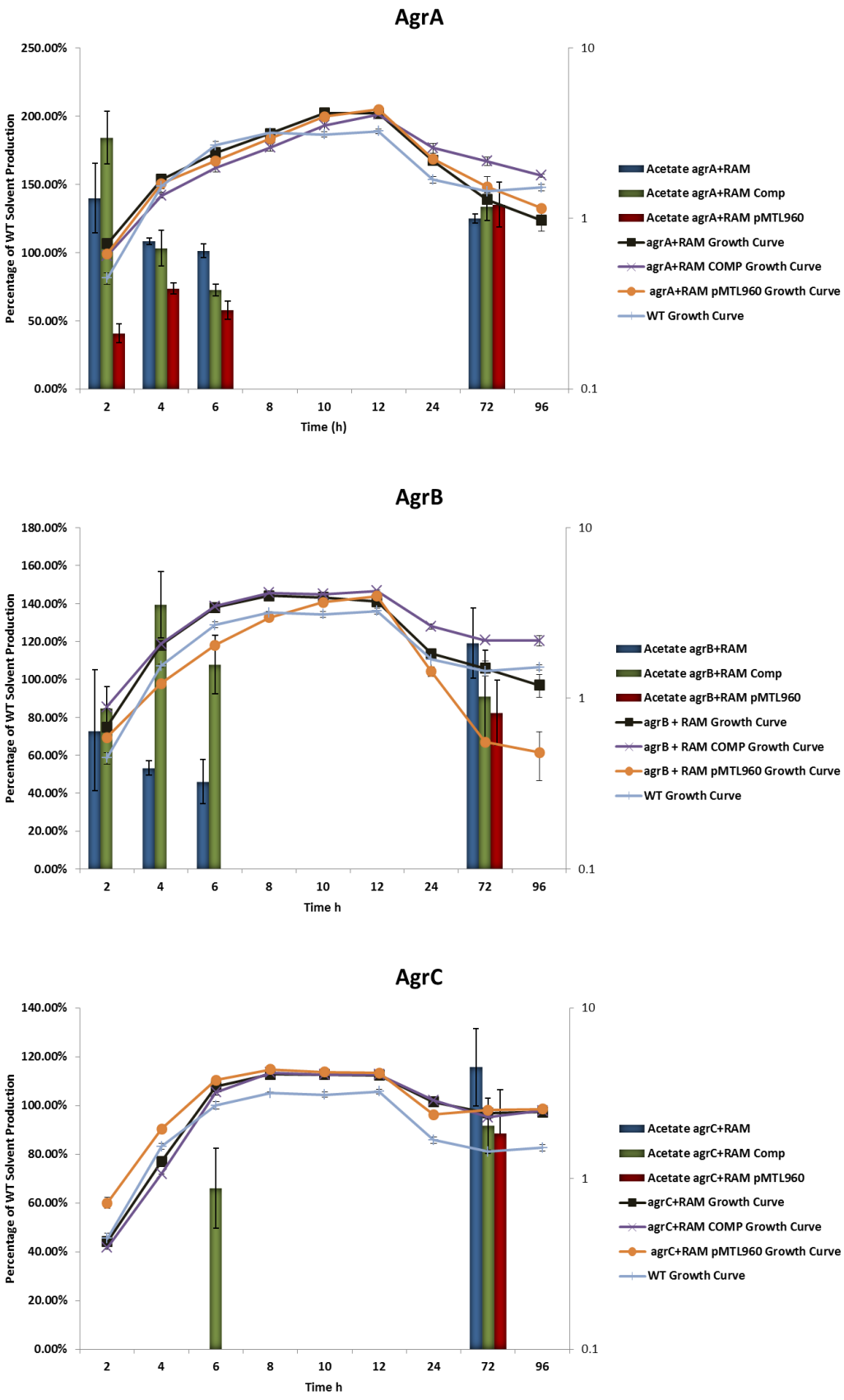


Figure 4.9 Graphs showing the acetate production of the three *agr* mutants with the RAM inserted as a percentage of WT production at the same stage of growth.

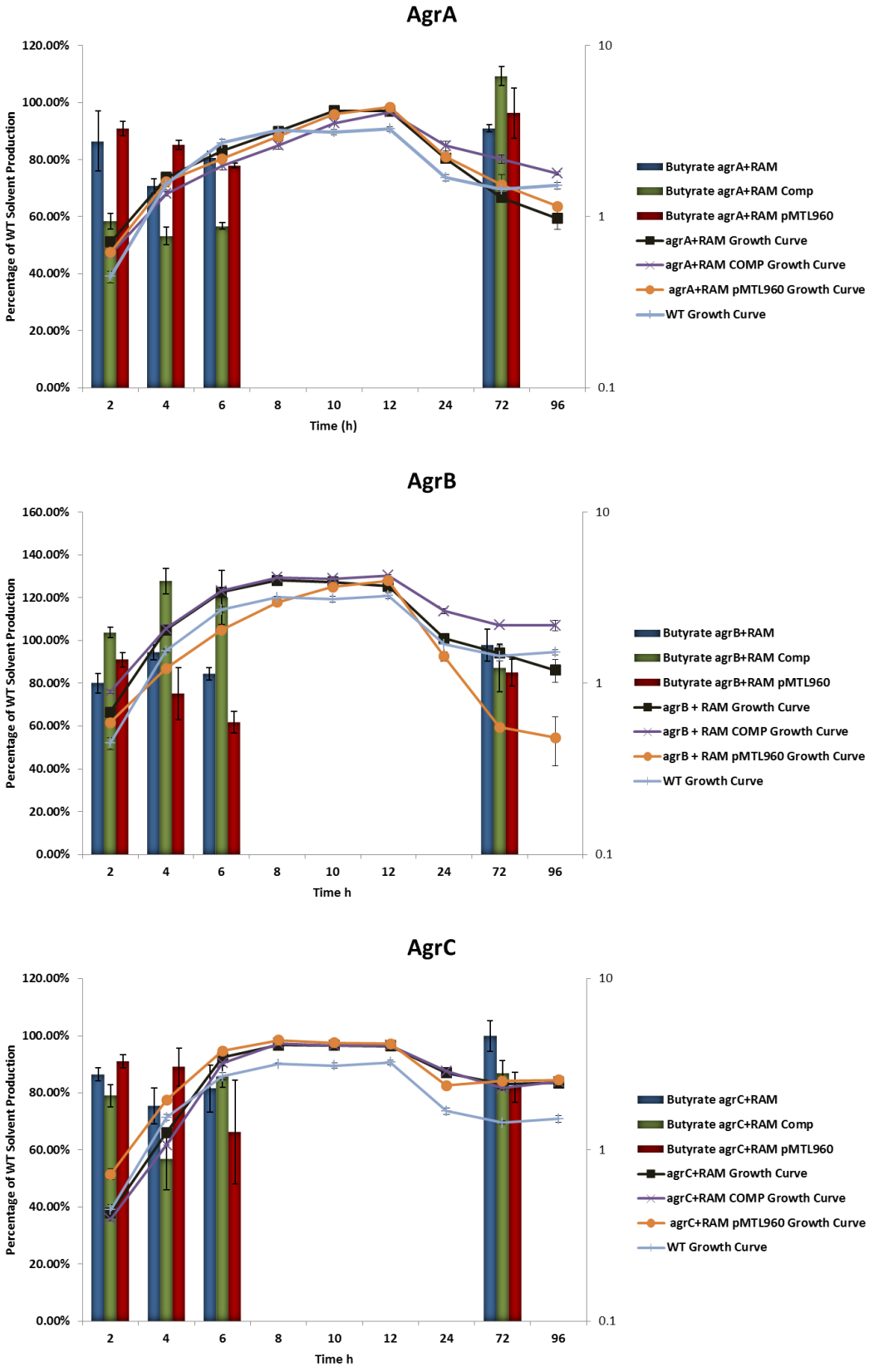


Figure 4.10 Graphs showing the butyrate production of the three *agr* mutants with the RAM inserted as a percentage of WT production at the same stage of growth.

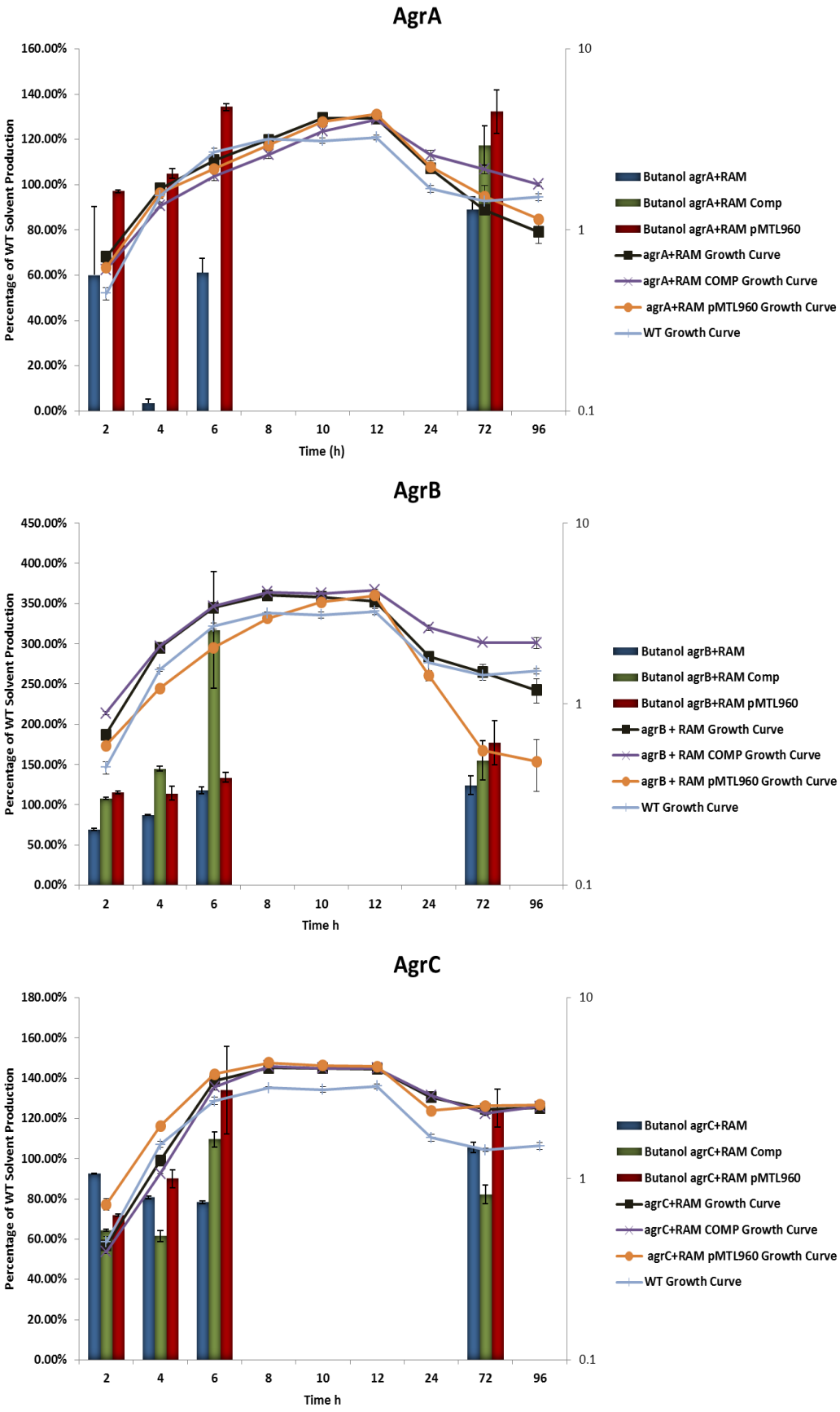


Figure 4.11 Graphs showing the butanol production of the three *agr* mutants with the RAM inserted as a percentage of WT production at the same stage of growth.

4.2.5 Removal of the selective marker using pMTL007C-E2

The RAM used in pMTL007 contains a very strong thiolase promoter in order to confer resistance from a single copy of the *ermB* gene. When the Targetron is inserted into a gene it is therefore a possibility that it may have polar effects on genes both downstream and upstream of the mutated gene. A mutant can be generated by insertion of the targetron in either a sense or antisense direction in respect to the target gene. The *agrC* mutant was generated by inserting the targetron in a sense direction. This means there is a possibility that the RAM's promoter will express in an antisense orientation which could interfere with the expression of genes upstream of the mutation, possibly through antisense RNA inhibition. The two other mutants were generated by inserting the targetron in the antisense direction where the promoter will promote transcription in a sense orientation. This could potentially enhance the expression of downstream genes.

These potential problems could be overcome if it was possible to excise the RAM with its strong promoter from the targetron after integration thereby minimising potential polar effects. John Heap achieved this by designing a new targetron vector with has FLP-sites flanking the entire RAM. This vector was designated pMTL007C-E2 (Heap *et al*., 2010). Three more retargeted vectors were made using pMTL007C-E2 which were targeted to the same three *agr* genes as the pMTL007 vectors. These three vectors were named pMTL007C-E2::Cac-agrA-141a, pMTL007C-E2::Cac-agrB-385a and pMTL007C-E2::Cac-agrC-121s which were targeted to *agrA*, *agrB* and *agrC* respectively. The same protocol was followed as described in section 2.16, to achieve integrants.

Three mutants were obtained as previously with the pMTL007 vectors and confirmed with PCR (Figure 4.12). The next step was to remove the *ermB* gene using FLP recombinase. Accordingly, pCLF-1 was transformed into the three mutants
using the method described in section 2.10, on average 75% of thiamphenicol resistant clones obtained were found to be Erm^S. To

confirm that the *ermB* gene had been excised, DNA from 10 putative clones was prepared and used with the appropriate primers in a PCR assay. In 100% of cases the product was consistent with removal of the *ermB* gene. As a final proof, the PCR product generated was subjected to nucleotide sequencing. This confirmed that the *ermB* gene had been excised entirely as expected (Heap et al., 2010). The 3 ClosTron-derived mutants in which the *ermB* gene had been subsequently removed together with the *agrA* mutant produced by allelic exchange ($agrA_{WD}$) (see section 4.2.2) were analysed for solvent production over the course of growth in batch culture. Due to the fact that complementation did not have any reproducible effect on the previous mutants it was decided not to transform the complementation vector or empty vector into the conventional *agrA* mutant or the new *ermB* deletion mutants.

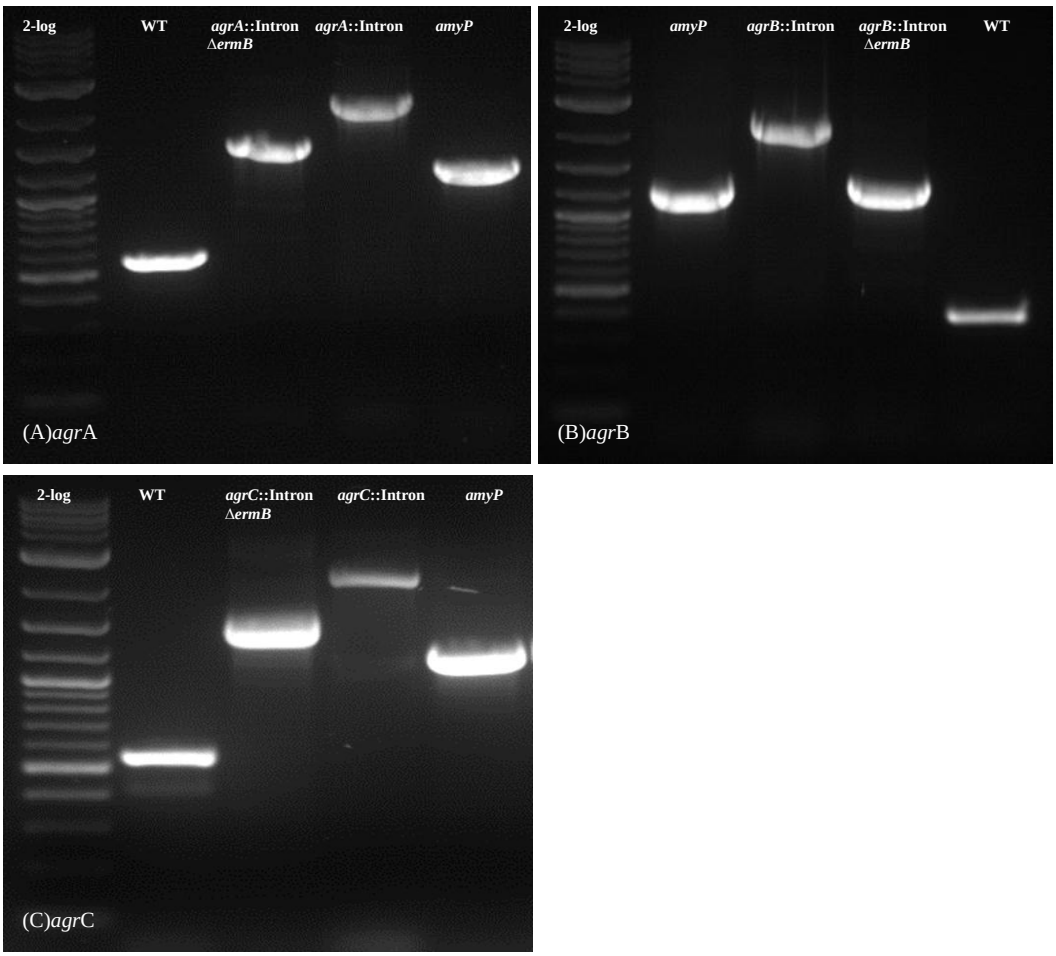


Figure 4.12 Characterisation of the three ClosTron mutants without *ermB*. (A) Primers SEQAGRAPF and SEQAGRAPR flank the site of insertion after base 141 of the *agrA* ORF.(B) Primers SEQAGRBPF and SEQAGRBPR flank the site of insertion after base 385 of the *agrB* ORF. (C) Primers SEQAGRCPF and SEQAGRCPR flank the site of insertion after base 121 of the agrC ORF.Each PCR used genomic DNA template isolated from the wild-type, the mutant in which the relative gene was inactivated by insertion of the intron from pMTL007C-E2, and finally the strain in which the ermB marker from this intron was removed by FLP recombinase. Primers CACampF and CAC amyPR were used with a genomic prep of each mutant to ensure pSOL-1 was present. Each *erm* closTron mutant was then sequenced with primers flanking closer to the insertion. The primers used were Cac-agrA-sF1 and Cac-agrA-sR1 for the *agrA* mutant, Cac-agrB-sF1 and Cac-agrB-sR1for the *agrB* mutant and Cac-agrC-sF1 and Cac-agrC-sR1for the *agrC* mutant.

4.2.6 Effect of *ermB* removal on *agr* knockout mutants

The same gas chromatography method described in section 2.17 which was used on the original ClosTron mutants containing the RAM was used to quantify the levels of product formation for each *agr* ClosTron mutant with the excised RAM and the *agrA* whole deletion mutant. The levels of product formation were expressed as a percentage of the same wild type levels as the mutants with a RAM. As with the original ClosTron mutants, the levels of product formation are shown at the first three time points

corresponding to exponential growth where acidogenesis occurs. The second time point is at 72 h where you would expect solvents to be in the supernatant as a result of solventogenesis.

4.2.6.1 Ethanol production

Figure 4.13 shows the levels of ethanol produced by the four mutants as a percentage of the wild type levels. The *agrB* and *agrA* ClosTron mutants showed lower levels of ethanol than the wild type levels at the earlier stages of growth. The *agrC* ClosTron mutant and the *agrA* deletion mutant showed higher levels of ethanol than the wild type levels at the earlier stages of growth. The *z*-scores for all 4 strains was 1.9640 at 2 and 4 h of growth. At 72 h of growth the *agrB* ClosTron mutant and the *agrA* deletion mutant showed levels of ethanol similar to the wild type levels (z = 1.525 and 0.282 respectively). At the same stage of growth the *agrC* and *agrA* ClosTron mutants showed levels of ethanol higher than the wild type strain (for both strains z = 1.964).

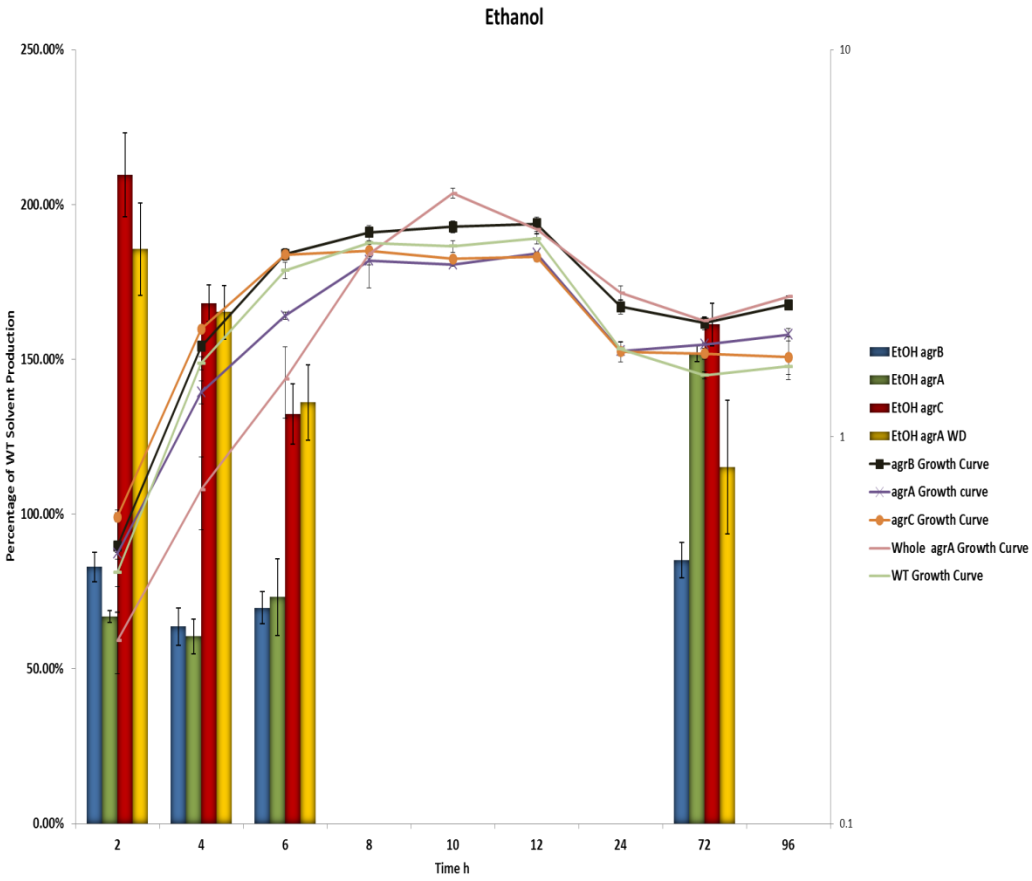


Figure 4.13 Ethanol production as a percentage of wild type production in the *agrA*, *agrB*, *agrC* and *agrA* (WD) mutants.

4.2.6.2 Acetone production

Figure 4.14 shows the levels of acetone production for each of the four mutants compared to wild type levels. At the early and the later stages of growth the *agrB* ClosTron mutant showed slightly lower levels of acetone production than the wild type strain. The *agrA* ClosTron mutant showed slightly lower levels of acetone production than the wild type strain at the early stages of growth but at 72 h there was less than 10% of the acetone present than the wild type strain produced. The *agrC* ClosTron mutant produced less than 10% of the wild type quantities of acetone throughout in both the early and later phases of growth. This was a different result to be obtained with the *agrC* mutant containing the RAM which produced higher levels of acetone than the wild type strain.

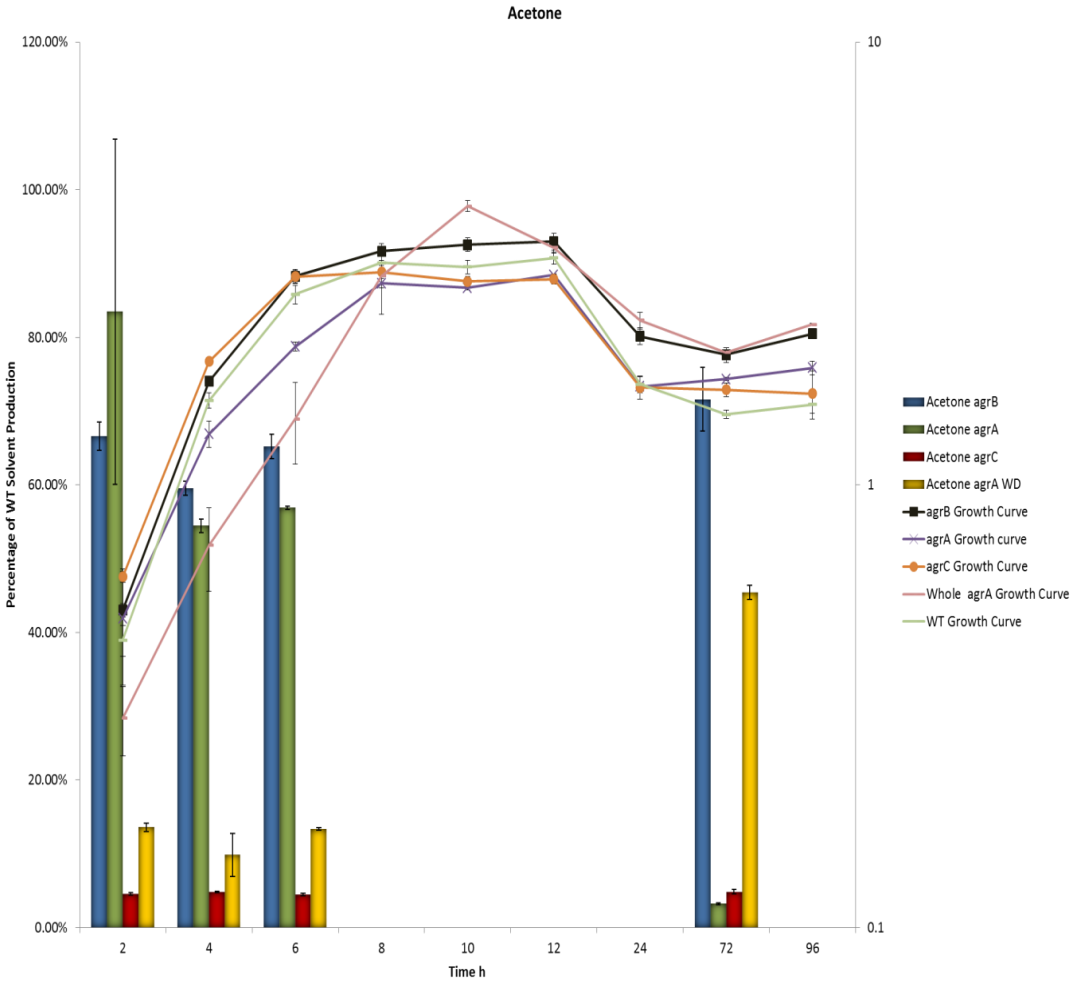


Figure 4.14 Acetone production as a percentage of wild type production in the *agrA*, *agrB*, *agrC* ClosTron and *agrA* deletion (WD) mutants.

The *agrA* deletion mutant showed less than 20% of the wild type quantities of acetone throughout the early phase of growth. At 72 h of growth the levels of acetone had increased to just below 50% of the levels of the wild type strain. These results are clearly different from those found with the ClosTron mutants containing the RAM as the level of acetone produced was higher than the wild type acetone levels throughout the growth curve. All four strains had a *z*-score of 1.964 except the *agrA* ClosTron mutant at 2 h of growth.

4.2.6.3 Acetate production

Figure 4.15 shows the levels of acetate production by the four mutants when compared to the wild type strain. The *agrB* ClosTron mutant showed lower levels of acetate production than the wild type strain at the earlier stages of growth to the wild type strain but showed similar levels of acetate at 72 h of growth. The *z*-score was lower than 1.960 at 4, 6 and 72 h of growth.
The *agrA* ClosTron mutant showed a similar pattern of acetate production to the *agrB* ClosTron mutant at the earlier stage of growth but at 72 h the levels of acetate were higher than the wild type (z = 1.964 at 6h and 72h).

The opposite was seen with the *agrA* whole deletion mutant which had higher levels of acetate at the earlier stage of growth than the wild type strain but lower levels at 72h (z = 1.964 at 4h, 6h and 72h).

The *agrC* Clostron mutant had higher levels of acetate present than the wild type strain at both the early and later stages of growth (z =1.964 at 2 h, 4 h and 72 h).

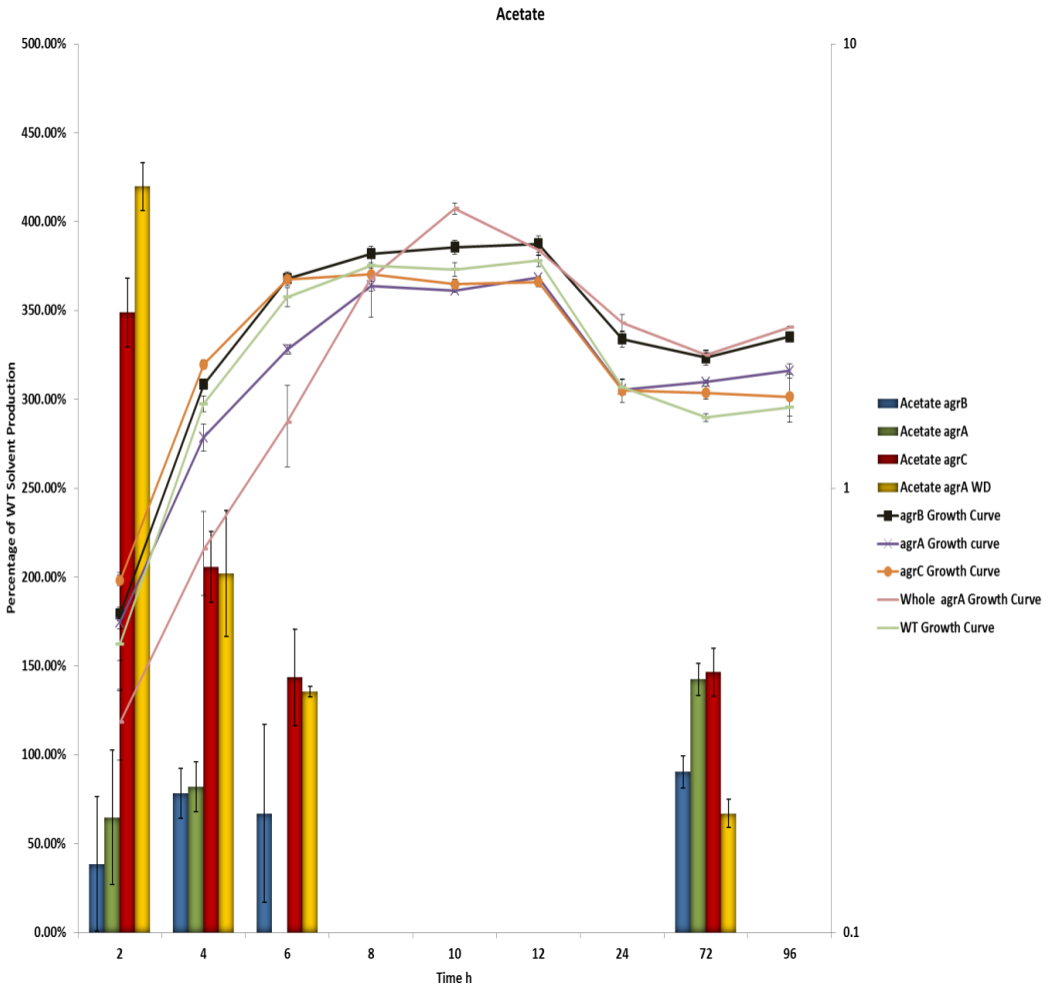


Figure 4.15 Acetate production as a percentage of wild type production in the *agrA*, *agrB*, *agrC* and *agrA* (WD) mutants.

4.2.6.4 Butyrate production

Figure 4.16 shows the levels of butyrate production by the 4 mutants when compared to the wild type strain. The *agrB* ClosTron mutant showed similar levels of butyrate production as the wild type strain at both the early and later stages of growth. The *z*-score was found to be less than 1.960 at all stages of growth and there was a high amount of variability indicating that any difference between the *agrB* mutant and wild type was not significant.

The *agrA* ClosTron mutant showed a similar pattern to the *agrB* ClosTron mutant in the early and late stages of growth but showed more of a decline

in levels over the early stages of growth. The *z*-score was found to be above 1.960 at 6h.

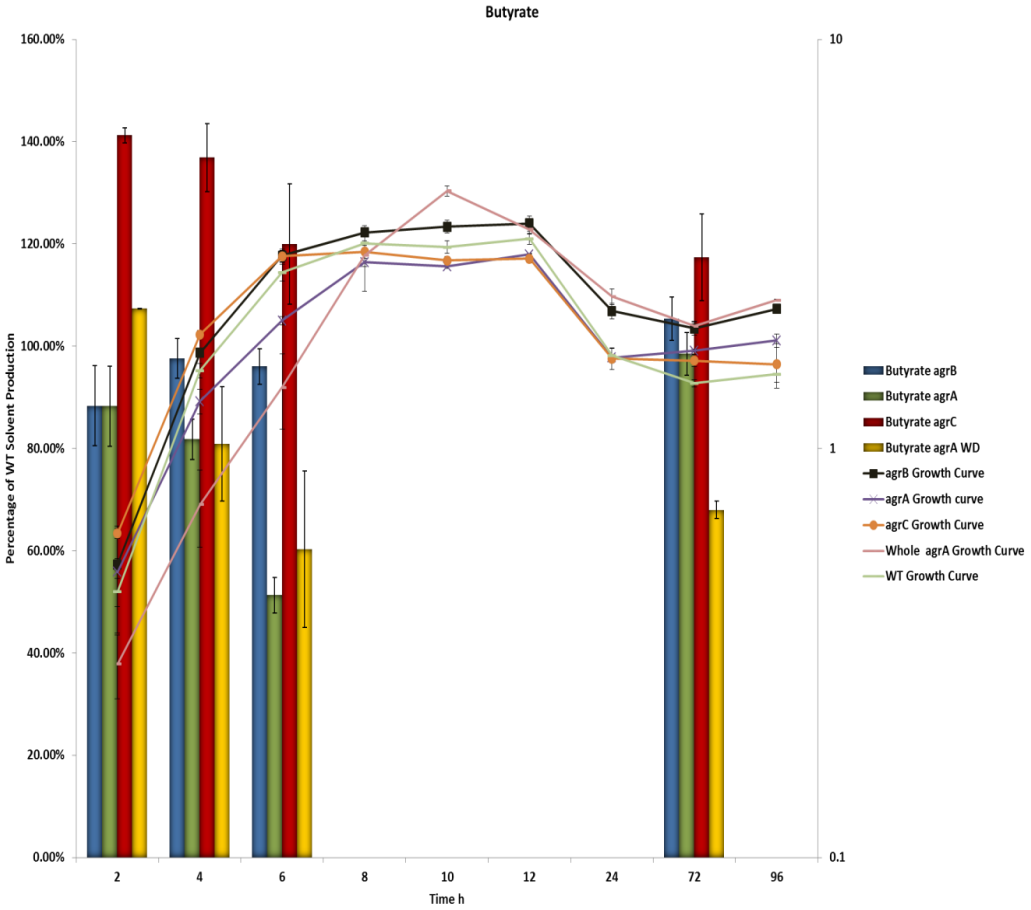


Figure 4.16 Butyrate production as a percentage of wild type production in the *agrA*, *agrB*, *agrC* and *agrA* (WD) mutants.

The *agrC* ClosTron mutant produced higher levels of butyrate than the wild type strain at all stages of growth although the *z*-score was found to be above 1.960 at 2h and 4 h only.

The *agrA* deletion mutant had higher levels of butyrate than the wild type at 2 h (z = 1.964) and showed a decline in levels of butyrate over the early stage of growth. It showed lower levels of butyrate than the wild type strain at 72 h of growth (z = 1.964).

4.2.6.5 Butanol production

Figure 4.17 shows the levels of butanol production by the 4 mutants when compared to the wild type strain. The *agrB* ClosTron mutant showed slightly higher levels of butanol as the wild type strain at the early stages of growth (z = 1.964). At 72 h growth the agrB mutants butanol levels were similar to the wild type strain (z = 0.282).

The *agrA* ClosTron mutant had slightly higher levels of butanol than the wild type at 2 h (z = 1.964) showed similar levels of butanol to the wild type strain at all the other time points shown (z =<1.960).

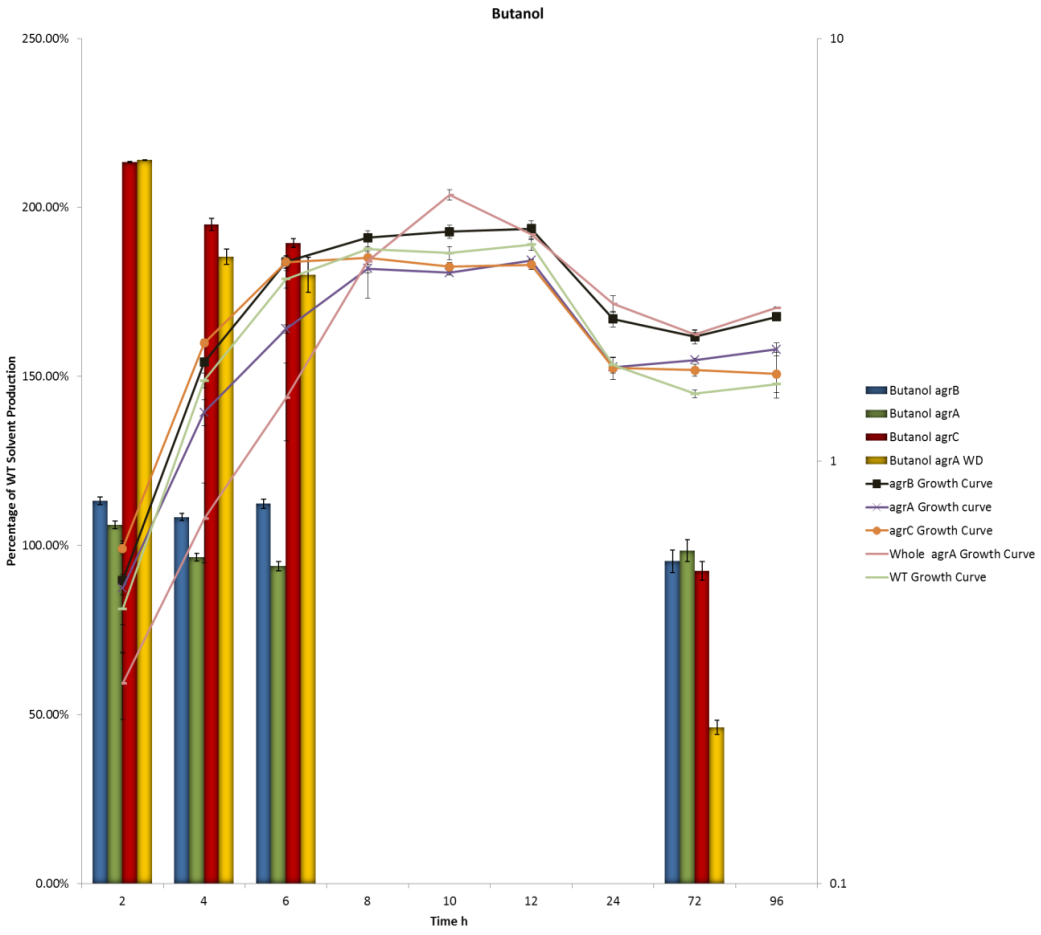


Figure 4.17 Butanol production as a percentage of wild type production in the *agrA*, *agrB*, *agrC* and whole deletion *agrA* mutants.

The *agrC* ClosTron mutant produced higher levels of butanol than the wild type strain at the early stages of growth but had a low z-score at 6 h

(1.091) but had low variability. This is probably linked to the variability in the wild type samples. The levels of butanol were similar to wild type at 72 h (z = 1.525).

The *agrA* whole deletion mutant followed a similar pattern to the *agrC* ClosTron mutant at the earlier stages of growth but had 50% of wild type the butanol levels at 72 h (z=1.964).

4.2.7 Sporulation

The development of a reproducible spore assay for measuring the rates of sporulation in *C.acetobutylicum* was undertaken given the demonstration that *agr* mutants were affected in sporulation in *C.botulinum* (Cooksley et al., 2010). It was thought highly likely that the *C.acetobutylicum* antisense mutants could be similarly affected.

The initial assay was based on a basic method in which the cells were grown in batch culture for 5 days (see section 2.18.1). This would give the cells enough time to sporulate. A sample of the culture was then heat treated or mixed with chloroform, serially diluted and then plated out on solid media. This was repeated with an untreated sample of the same culture. The resulting colonies were counted and the number of treated colonies was compared with the number of untreated colonies. This method should in theory, give the number of spores produced compared to the number of viable cells.

A number of problems were encountered in the development of this assay. The cells did not seem to sporulate consistently in the CBM media used. This could have been due to ‘acid crash’ caused because of insufficient buffering of the medium. There were also problems diluting the cultures in order to get the appropriate number of colonies for each dilution. This was probably due to the clumping of *C.acetobutylicum* in liquid media. The assays were repeated using chloroform to treat the cells instead of heat treatment and this made the results more consistent although a liquid

based assay was unlikely to produce consistent enough results to be useful. This is reflected in the literature as although sporulation has been studied in *C.acetobutylicum* for many years there have been no reports of this type of assay being used.

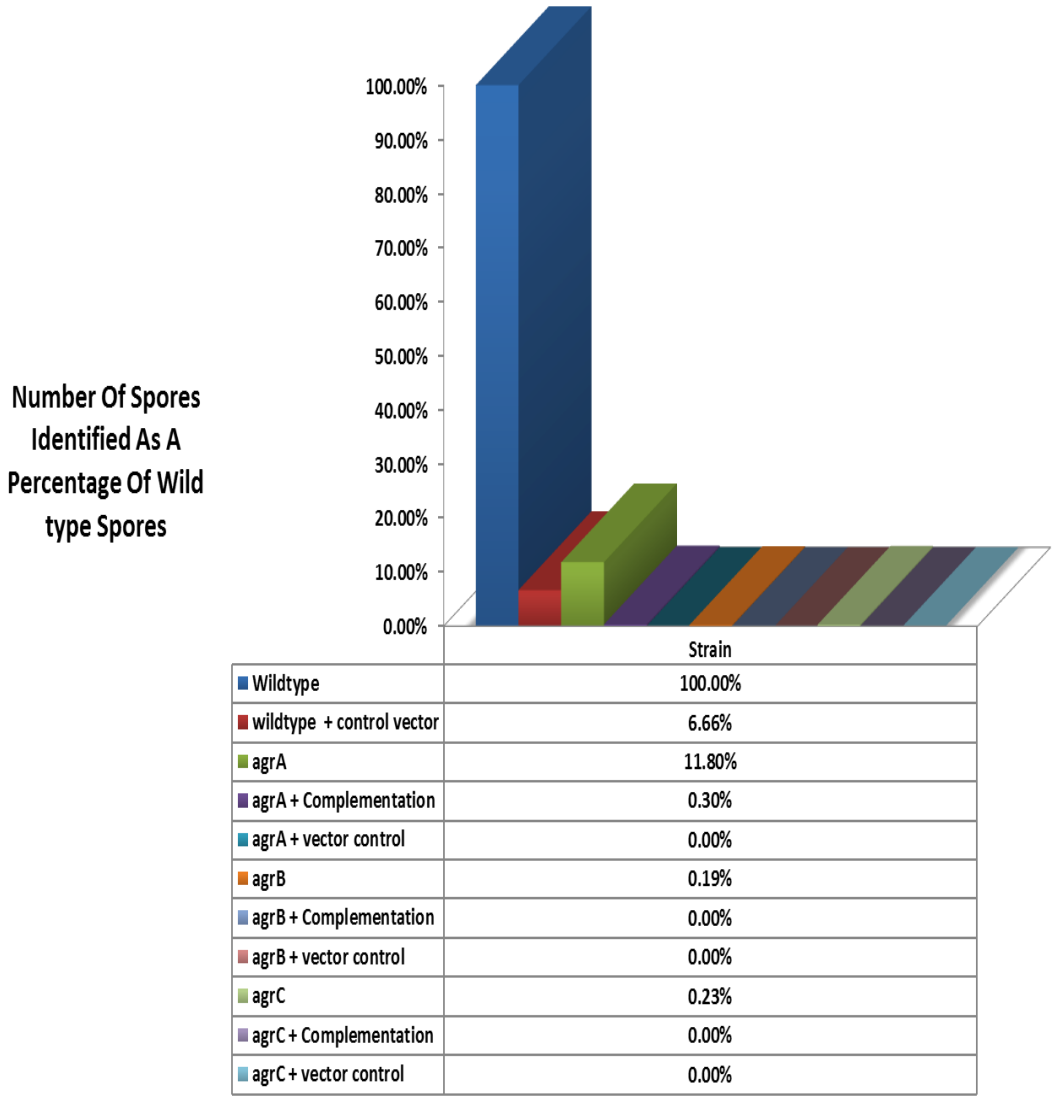
The assays performed although not ideal did not show any significant differences between the control and the antisense strains which had the pJS06 and pJS07 vectors with respect to sporulation. It was therefore difficult to know whether it was because the vectors were not repressing *agr* translation or if Agr had no effect of sporulation. Without antibodies to Agr proteins it was not possible to do a western blot to see an antisense effect and this was not used at this stage due to future developments.

In order to potentially improve any possible effects of the antisense RNA, two more vectors pJS08 and pJS09 were made targeting *agrBD* and *agrCA* respectively using the method described in figure 4.1. These would each produce longer antisense transcripts. These were effectively transferred into *C.acetobutylicum* but it was decided to halt the use of antisense RNA as by this stage, Clostron technology had become available.

4.2.7.1 Effect of *agr* knockouts on sporulation

After examining the effect of *agr* gene knockouts on product formation, the next step was to investigate the effect of *agr* knockout on sporulation. No phenotypic assay to measure sporulation in *C.acetobutylicum* has been published to date and as previously described, many problems were encountered during the antisense RNA studies. The main reason for the problems in the development of a quantitative assay is the fact that *C.acetobutylicum* does not sporulate very well, or at least consistently in liquid culture. *C.acetobutylicum* does sporulate well on solid media, and this provided the basis for a viable assay. As described in section 2.18.2, it relies on leaving the cells to sporulate on CBM media for 5 days. A colony containing cells and potential spores was streaked off the medium and

diluted 1:20 with water. A 5 µl sample was then placed on a microslide and 3 random photographs were taken at 400 x magnification and the number of vegetative cells and spores were counted manually. This was repeated twice for each plate resulting in 3 photographs taken for each of the three colonies for each strain. The number of spores was then recorded as a percentage of the total number of vegetative cells. Examples of these slides are shown in figure 4.19.



	Strain
Wildtype	100.00%
wildtype + control vector	6.66%
agrA	11.80%
agrA + Complementation	0.30%
agrA + vector control	0.00%
agrB	0.19%
agrB + Complementation	0.00%
agrB + vector control	0.00%
agrC	0.23%
agrC + Complementation	0.00%
agrC + vector control	0.00%

Figure 4.18 Graph showing the average number of spores for each of the three ClosTron mutant strains containing the RAM as a percentage of the total number of wild type spores observed. The counts for the complementation and plasmid control strains are also shown. The spores and cells were counted from 3 random pictures at 400 x magnification taken from 3 separate colonies giving an average number from 9 separate pictures.

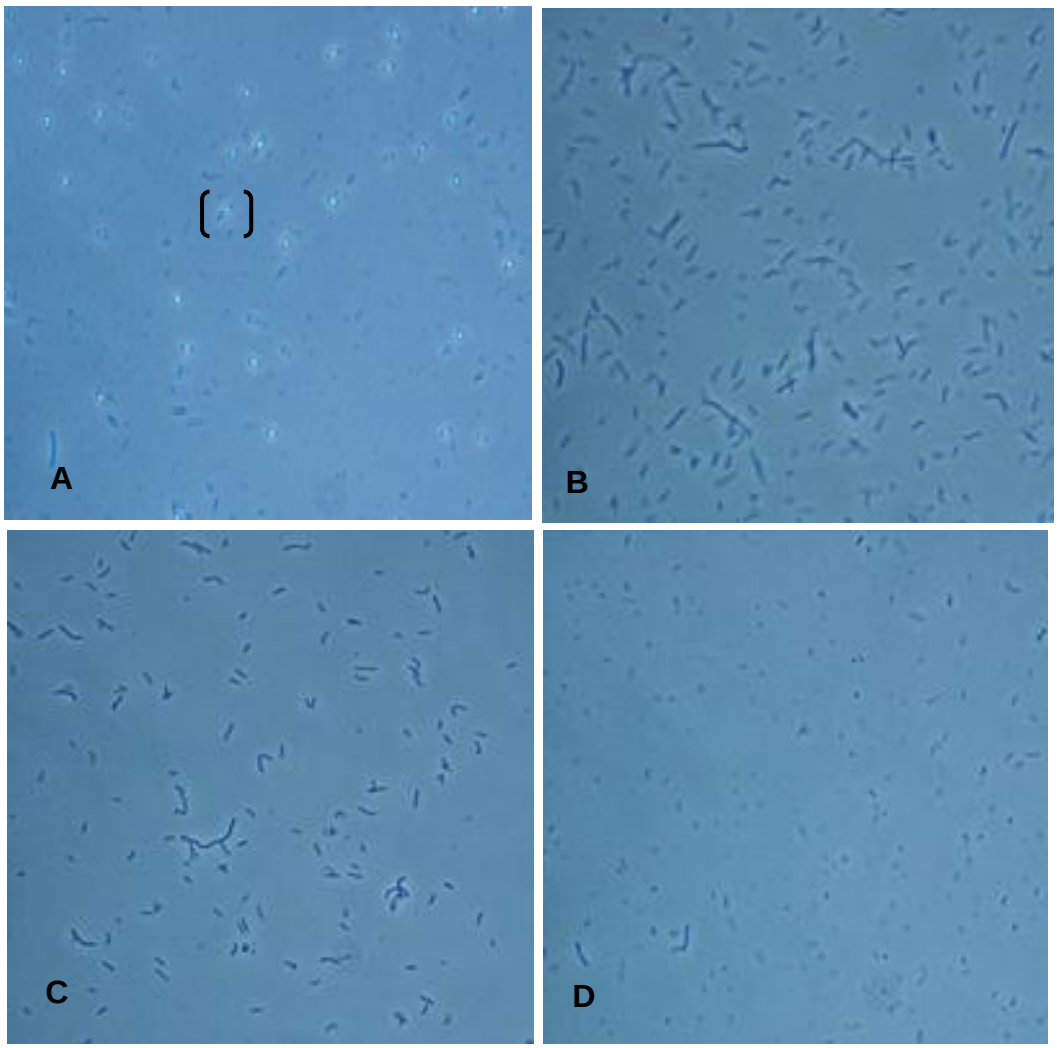


Figure 4.19. Four examples of the pictures taken for the solid medium based spore assay described in section 2.18.2. Picture A shows the wild type *C.acetobutylicum* 824 strain after 5 days growth. One of the spores is highlighted with brackets. Pictures B, C and D show the *agrA*, *agrB* and *agrC* mutants respectively. As can be seen there are no spores present in these pictures.

The pictures taken of the three *agr* mutants containing the *ermB* gene showed a reduced number of spores when compared to the numbers shown with the wild type strain. The *agrB* strain showed an increased level of sporulation as compared to the two other mutants but this was still reduced when compared to the wild type strain.

As found with the GC results the complementation vectors did not effectively complement the differences between the mutant and wild type strains. The strain containing the empty vector pMTL960 showed greatly reduced levels of sporulation when compared to wild type. This suggested

it was the presence of the vector itself which was affecting sporulation rather than any effects from over expression of the *agr* genes within the plasmids.

The sporulation assay was repeated for the *ermB* mutants. As can be seen in figure 4.19, these three mutants produced few or no spores under the same conditions as wild type. Some spores could be seen with the *agrB* mutant as for the mutant containing the RAM but the number was significantly lower than the wild type strain.

4.3 Discussion

Much of what is described in this chapter shows evidence of the rapid advancement in the development of molecular tools for use in the study of *C.acetobutylicum*. The initial development of antisense RNA for gene disruption was overtaken by new developments in gene knockout technology, the ClosTron. The time taken to obtain a single mutant using double crossover (over 6 months) reflects how difficult it has been to obtain gene disruption with this organism. This emphasises what a breakthrough the ClosTron system is, enabling gene disruption in a matter of a few weeks rather than months.

The two most intensively studied phenomena in *C.acetobutylicum* are its metabolic pathways and sporulation. For this reason once the *agr* mutants were obtained it was decided to look at the effect of *agr* knockout on both product formation and sporulation. This was hampered by the lack of a definitive sporulation assay and a suitable protocol for solvent analysis which did not rely on specialised apparatus.

A Gas Chromatograph had to be purchased specifically to analyse the products of *C.acetobutylicum* and the comparison of product formation between the wild type strain and the *agr* mutants was the first to be conducted with the new machine.

The tendency for *C.acetobutylicum* to sporulate readily on solid media was used to advantage in the development of a very simple assay. This assay gave very consistent results when compared to a liquid based assay where *C.acetobutylicum* has a tendency to sporulate unpredictably. The only problem is the fact the spores have to be counted subjectively on digital photographs. There was, however a difference between the numbers of spores in the wild type strain and those of the mutant strains (most photos in figure 4.19 showed no spores at all). The number of photos which had to be counted made the process very time consuming and for this reason a reliable liquid based assay would be desirable.

The results presented in this chapter show some evidence that knocking out the putative *agr* genes does have a phenotypic effect on the organism although the sample numbers were low. There was only one sample from three batch cultures taken for each mutant and the complemented strains used. The standard error of the mean was used to determine the error bars in the graphs presented. In some cases the variability was too great to determine a significant difference between wild type, mutants and control strains. The Mann-Whitley *U*-test was used to obtain a *z*-score for comparing the levels of product formation between the mutant and wild type strains. The tables showing these *z*-scores are in Appendices I and II. A sample size of three is the smallest which can be used for this test and it is necessary to assume that the samples are part of a normal distribution. The fact that no *z*-score was found to be above 1.964 suggests that larger sample sizes are needed to provide more statistically robust evidence for the differences observed in product formation between the *agr* mutants and the wild type control. Mutations in each of the genes with the original ClosTron, in which the intron insertion of the mutant generated additionally contained a copy of the *ermB* gene, showed some effects when it came to solvent production. The strains which contained the complementation plasmid and the empty vector control showed inconsistent results with levels of solvent production in all 3 *agr* mutants. The complementation vectors did not show a consistent pattern of complementation in the 3 strains.

There have been other reports of plasmids in *C.acetobutylicum* affecting fermentative growth (Walter *et al.,* 1994; Tomas *et al.,* 2003a). It was for this reason that this vector and the control vector were not used to complement the mutants without the presence of the RAM. Overall the *agrA* mutant and the *agrB* mutant containing the *ermB* gene did not have an effect on product formation over the growth cycle. The levels of most of the products produced by these strains did not differ a great deal from the wild type strain. When it was noted that there was higher or lower levels this could only be a 20% difference in the levels. The levels of acetate could not be calculated effectively for some of the earlier time points which is disappointing as this would be the time point in which acetate would be expected to be produced. It should be noted that the levels of acetate were not very different from those of the wild type at the later stage of growth for the three mutants. The *agrC* mutant with the *ermB* gene did show some unusual levels of acetone production with product level being acutely higher at 72 h with the *agrC* mutant when compared to the wild type strain. Because the *ermB* gene in the *agrC* mutant is inserted in the sense orientation there is the possibility of the *thl* promoter having a positive effect on downstream genes. Whether it is this phenomenon which is causing the increased levels of acetone production is hard to confirm, although this overproduction is not seen with the *agrC* mutant without the *ermB* gene.

The *agr* mutants lacking the *ermB* gene showed different results to the mutants containing the *ermB* gene which could be attributed to the presence of the *ermB* gene itself. Therefore the results obtained with the mutants minus the *ermB* gene could be regarded as a more reliable interpretation of the impact of *agr* on product formation.

In terms of ethanol, the *agrA* deletion and *agrC* ClosTron mutants did show higher levels of product formation than the *agrA* and *agrB* ClosTron mutants, when compared to the wild type control at the earlier stage of growth ($z = 1.964$). At the later stage of growth all the mutants except the

agrB mutant showed higher levels of ethanol then the wild type. This is the stage were ethanol would be present in higher quantities as a result in solventogenesis. With acetone production the levels of product were small in all strains without the selective marker. The difference was particularly large with the *agrC* and *agrA* ClosTron mutants which showed much lower levels of acetone production when compared to wild type. As previously described, this was not observed with the ClosTron mutants containing the RAM. The amount of acetate produced by the mutants without a selective marker followed a similar pattern to ethanol production .The amount of butyrate produced by the mutants did not seem to be different from the wild type levels, although the levels of butyrate were lowest in the *agrA* whole deletion strain. The levels of butanol for all three ClosTron mutants without the RAM was similar to the wild type levels at 72 h when levels of this solvent would be expected to be higher than the earlier stages of growth. The *agrA* deletion mutant showed slightly lower levels at this stage.

Interestingly the *agrA* deletion mutant showed different levels of product formation than the *agrA* mutant with *ermB*. In the ClosTron derived mutants the *agr* genes are disrupted while in the other mutant using the Soucaille laboratory methods the whole gene is excised. This may explain the difference between the two mutants.

A major issue with these assays is the fact that the 7 mutants and the wild type strain do not follow the published pattern of product formation or the levels of product formation (Mermelstein and Papoutsakis 1993). A clear transition between acidogenic and solventogenic phases would be expected as the cells reach the stationary phase of growth. There was clearly a low amount of product formed from these batch cultures. The ratio of solvents produced was also different from published accounts of 3:6:1 for acetone, butanol and ethanol respectively. The ratios of solvents detected by the gas chromatograph were 8:15:13 for acetone, butanol and ethanol respectively. A controlled pH 5.5 batch culture of *C.acetobutylicum* has been shown to produce a maximum ~155mM of acetate and butyrate

at 46h and 16 h growth respectively (Mermelstein and Papoutsakis 1993). The same study showed a maximum production of ~ 130 mM butanol, and ~80 mM acetone at 46 h growth and ~20 mM of ethanol at 30 h (Mermelstein and Papoutsakis 1993). The wild type control used in this study produced a peak of 56 mM of acetate at 24 h growth and 86 mM of butyrate at 96 h growth. The maximum production of each solvent produces was 13 mM of ethanol, 8 mM of acetone and 15 mM of butanol at 24h of growth. Most laboratories use buffered batch cultures which are grown in dedicated bioreactors or continuous cultures. The medium in which the cells where grown was based on a formulation which was passed on from the laboratory of Hubert Bahl in Rostock but it appears that in these conditions the organism does not produce high levels of solvent products.

A new medium has recently been developed by the *Clostridium* laboratory at Nottingham University in which the organism produces greater quantities of product but the assays could not be repeated with the new medium in the time given for this study. A recent study using ClosTron derived *agrB*, *agrC* and *agrA* mutants has found no effect of *agr* on solvent formation under the conditions tested (Steiner *et al*., 2012).

4.3.1Sporulation

The plate based spore assay clearly shows that knocking out the putative *agr* genes has a pronounced negative effect on sporulation. Many studies have shown that solventogenesis and sporulation are linked but not exclusively (Scotcher *et al.,* 2005). It is therefore possible that these *agr* homologues have an effect on sporulation but not on solventogenesis. Sporulation in *C.acetobutylicum* is poorly understood compared to organisms such as *B.subtilis* which does not have a homologue of *agr.* Spo0A the main global regulator of sporulation in *B.subtilis* and *C.acetobutylicum,* also controls solventogenesis in the latter organism. It is, therefore, interesting to find a putative global regulatory system which appears to effect sporulation only. Three *agr* mutants have been shown to

have a reduced number of spores produced in liquid and solid plate based assays (Steiner et al., 2012).

4.3.2 Complementation

The presence of pMTL960 and the control vectors based on this plasmid also had a negative effect on sporulation.The fact that the empty vector control had an effect suggests that it is the presence of the plasmid itself in the strain which is effecting sporulation. The vector control strain was grown with and without antibiotic which resulted in little difference in sporulation suggesting that the inhibition of sporulation was not directly the result of the addition of antibiotics. As mentioned previously there have been reports of plasmids interacting with their host to change growth and product formation (Walter *et al.,* 1994; Tomas *et al.,* 2003a; Scotcher *et al.,* 2005). Furthermore a control test was conducted to ensure the plasmid was not lost before sporulation. A recent study has managed to complement *agr* mutants using a diferent plasmid vector (Steiner *et al*., 2012).

It is unfortunate that the cells could not be complemented effectively as this would provide confirmation that it was the single deletion of the gene which is having the effect and not some other disruption. It remains unlikely however that there would be more than one insertion event with the ClosTron system. Unfortunately due to time constraints a Southern blot was not used to confirm there was only a single insertion in the mutants.

Overall the analysis of the effect of *agr* gene inactivation could not provide conclusive evidence that the *agr* mutants produced different amounts of acids and solvents than the wild type strain. There is some evidence to suggest that *agr* is involved in the regulation of sporulation. Further evidence for why the *agr* genes have this effect could be found by studying the *agr* genes' regulatory effects on the genome of *C.acetobutylicum*. This approach is described in the following chapter.

5.0 The effect of *agrB* knockout on gene expression in *C. acetobutylicum*.

5.1 Introduction

As illustrated in the previous chapter the gene knockouts resulted in strains which show a phenotypic difference to wild type strain, particularly in relation to sporulation. To investigate the impact of *agr* on the transcriptome, microarray analysis was undertaken.

Microarray analysis is a powerful tool with which to look at the effect of gene knockout on the expression of all other genes in an organism. This technique has been used successfully by the laboratory of Eleftherios Papoutsakis to look at the *C.acetobutylicum genome* in great detail (Tummala *et al.,* 2003b; Alsaker *et al.,* 2004; Alsaker and Papoutsakis, 2005; Borden and Papoutsakis, 2007; Parades *et al.,* 2007; Jones *et al.,* 2008).

The study by Alsaker and Papoutsakis (2005) specifically examined gene expression over a time course experiment using continuous culture. They also compared gene expression over a time course between wild type *C.acetobutylicum* and the M5 strain which lacks pSOL-1. Another significant study by the same laboratory (Jones *et al.,* 2008) investigated the sporulation specific transcriptional program over the stationary phase of growth. The batch culture was grown in a pH controlled bioreactor. These two papers will be referenced repeatedly in this chapter due to their similarity to this study.

The *agrB* mutant was chosen as a suitable candidate for microarray analysis as AgrB plays a pivotal role in the generation of the putative signal AIP molecule. The *agrB* knockout mutant which did not contain the RAM was used for the microarray study. The difference between the expression of genes in the *agrB* knockout mutant was compared to the equivalent gene expression in the wild type organism.

One of the major obstacles in a microarray study is obtaining RNA of sufficient quantity and quality. The method described in section 2.14.1 made it possible to get sufficient RNA of the desired quality for the microarray which was performed by Victoria Wright at the Centre for Biomolecular Sciences at the University of Nottingham as described in section 2.14.

The large number of differentially regulated genes which were found in this study makes it difficult to analyse the data efficiently. In order to make it easier the genes have been separated according to what orthologous proteins they encode.

Clusters of Orthologous Groups of proteins (COGs) were delineated in NCBI by comparing protein sequences encoded in complete genomes, representing major phylogenetic lineages. Each COG consists of individual proteins or groups of paralogs from at least 3 lineages and thus corresponds to an ancient conserved domain. All the genes found to be differentially regulated were sorted into their appropriate Cluster of Orthologous Group (COG) of proteins and Table 5.1 gives a description of each cluster. The large number of genes differentially regulated means it is beyond the scope of this thesis to discuss all the genes in detail. Because of the initial findings of this study the main focus will be on those genes found to be differentially regulated which are involved with solventogenesis, sporulation, carbon metabolism and motility. All the genes which were differentially regulated significantly (p value <0.01) and have a greater than two fold change in expression have been placed in appendix 1.

5.2 Results

The microarray analysis revealed 168 genes up-regulated in the *agrB* mutant at the earlier stage of growth when compared to the wild type strain. These genes could therefore be negatively regulated by the *agr* system in *C.acetobutylicum* during the exponential stage of growth. Figure

5.1 shows the highest number of genes was in COG group K which are transcription related genes. There were also 14 genes with a predicted function for carbohydrate transport and metabolism.

The number of genes in each COG group up-regulated in the wild type strain at the earlier stage of growth when compared to the *agrB* mutant is shown in figure 5.2. Overall there were 125 genes up-regulated at this stage of growth. Some of the COG groups shown to be over represented are involved in signal transduction systems, cell and envelope biogenesis and energy production.

At the later stage of growth corresponding to the transition phase of growth there were 220 genes up-regulated in the *agrB*

J	Translation, ribosomal structure and biogenesis
K	Transcription
D	Cell division and chromosome partitioning
L	DNA replication, recombination and repair
O	Post translational modification, protein turnover, chaperones
M	Cell envelope biogenesis, outer membrane
N	Motility & secretion
P	Inorganic ion transport and metabolism
T	Signal transduction mechanisms
C	Energy production and conversion
G	Carbohydrate transport and metabolism
E	Amino acid transport and metabolism
F	Nucleotide transport and metabolism
H	Coenzyme metabolism
I	Lipid metabolism
Q	Secondary metabolites biosynthesis, transport and catabolism
R	General function prediction only
S	Function unknown
V	Defence mechanisms

Table 5.1 The description of each COG group

mutant when compared to the wild type strain. The number of genes in each COG group is shown in Figure 5.3. Some of the over-represented groups include those associated with transcription, translation and post translational modifications.

The wild type strain had 174 genes up-regulated at the later stage of growth. Figure 5.4 shows the number of genes in each COG group at this stage of growth. Some of the over represented COG groups included those associated with translation, post translational modification, cell envelope biogenesis, carbohydrate transport and metabolism and cell division.

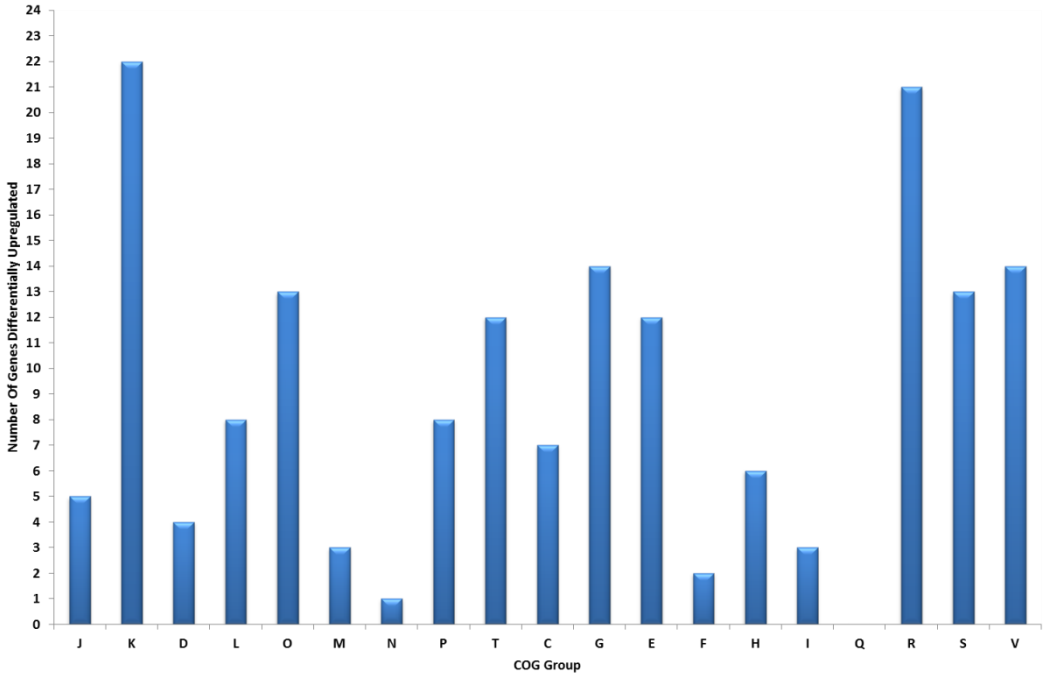


Figure 5.1 The number of genes in each COG group which are differentially up-regulated in the *agrB* mutant when compared to the wild type strain at the exponential stage of growth.

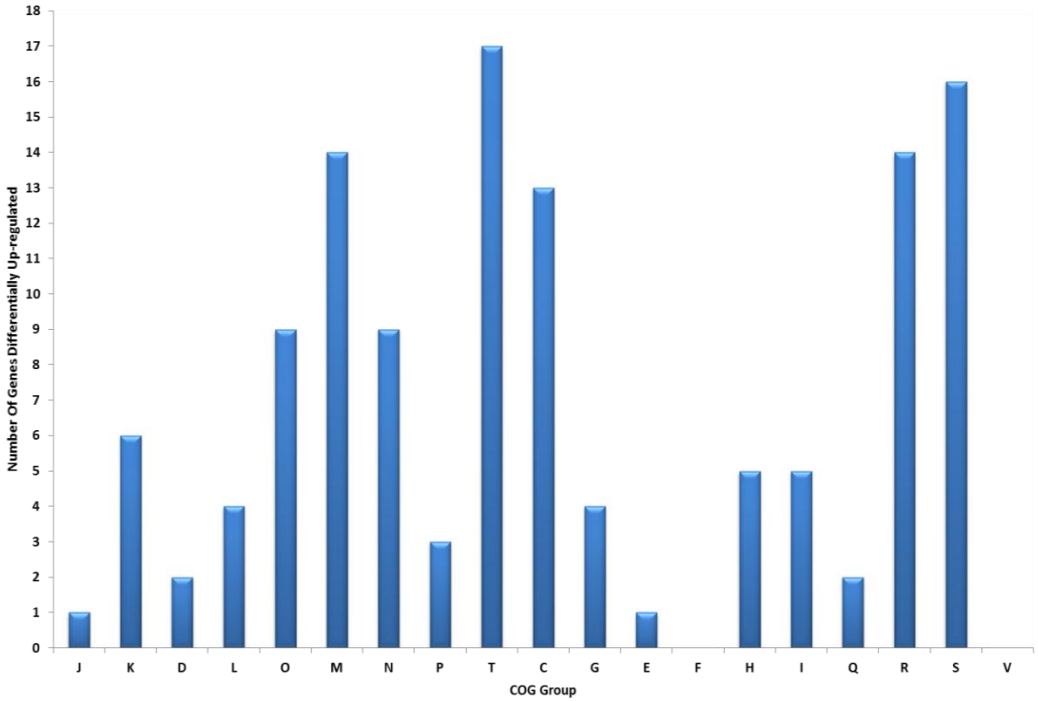


Figure 5.2 The number of genes in each COG group which are differentially up-regulated in the wild type strain when compared to the *agrB* mutant at the exponential stage of growth.

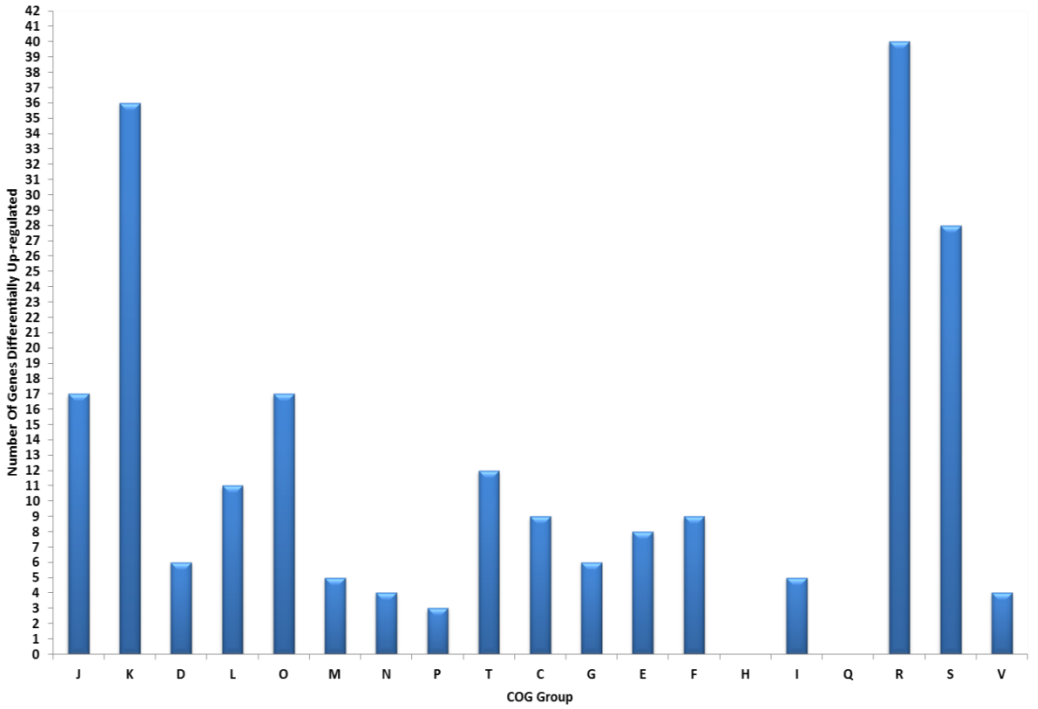


Figure 5.3 The number of genes in each COG group which are differentially up-regulated in the *agrB* mutant when compared to the wild type strain at the transitional stage of growth.

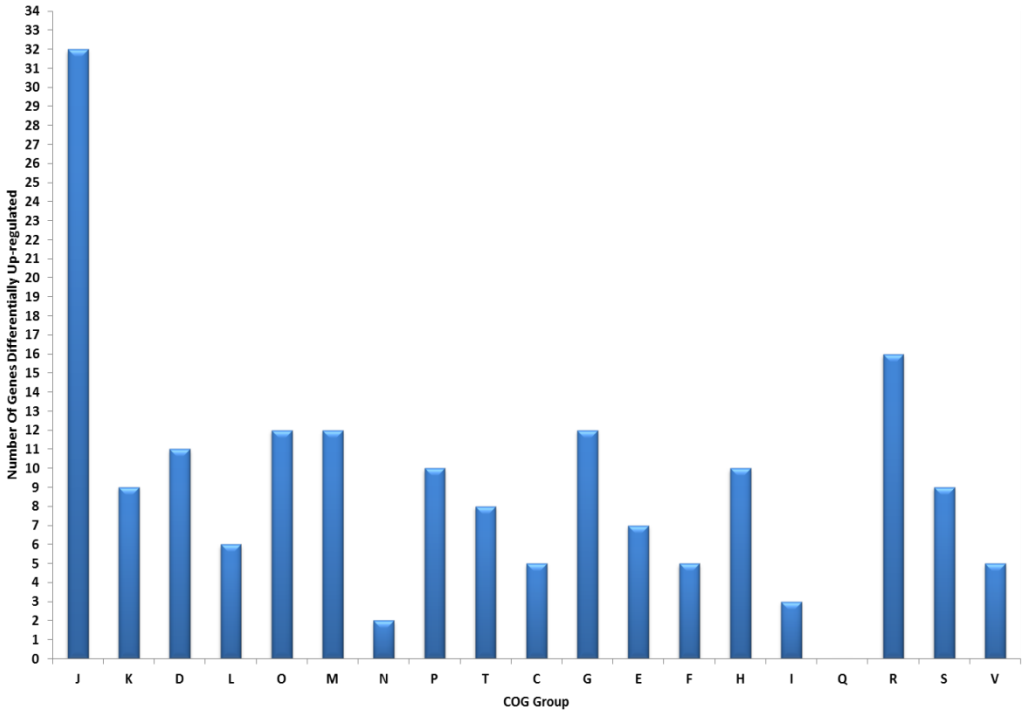


Figure 5.4 The number of genes in each COG group which are differentially up-regulated in the wild type strain when compared to the *agrB* mutant at the transitional stage of growth.

5.2.1 Genes Involved in Solventogenesis

A relatively small percentage of all the genes known to be involved in solventogenesis were found to be differentially regulated in an *agrB* dependent manner. The placement of these genes in the solventogenic pathways is shown in figure 5.5. CAC0980 designated as *pfl*B was found to be up-regulated 3-fold in the late stage *agrB* mutant as compared to the wild type at the same stage.

CAC0980 is involved in the conversion of pyruvate into acetyl-CoA and these results suggest it is negatively regulated by *agrB*. As can be seen in figure 5.5 this enzyme provides a feeder molecule for many of the reactions necessary for product formation. The pSOL-1 based gene, CAP0035 annotated as *adhE* was also found to be up-regulated 2.7 fold at the late stage of growth in the *agrB* mutant as compared to the wild type at the same stage. Interestingly this gene was found to be responsible for butanol production in alcohologenic cultures and *C.acetobutylicum* is thought to be unique in having two copies of this gene (Fontaine *et al.,* 2002). This gene was earlier found to be up-regulated at the transitional

phase between the log and stationary phases of growth (Jones *et al*., 2008).

CAC2012 and CAC2016, the two genes encoding enoyl-CoA hydratases annotated as *fadB* were up-regulated at the earlier stage of growth (2.074- and 3.807-fold respectively) in the wild type strain as compared to the *agrB* mutant at the same stage. CAC2009, a 3-hydroxy acyl-CoA dehydrogenase was up-regulated at the earlier stage of growth 5.2-fold in the wild type strain as compared to the *agrB* mutant at the same stage.

Another gene which seems to be positively regulated by *agr* was CAC3298, a NADH dependent butanol dehydrogenase B. Expression was found to be 2-fold higher in the wild type strain at the later stage of growth when compared to the *agrB* mutant at the same stage. No other genes implicated in product formation were found to be differentially regulated in an *agr* dependent manner.

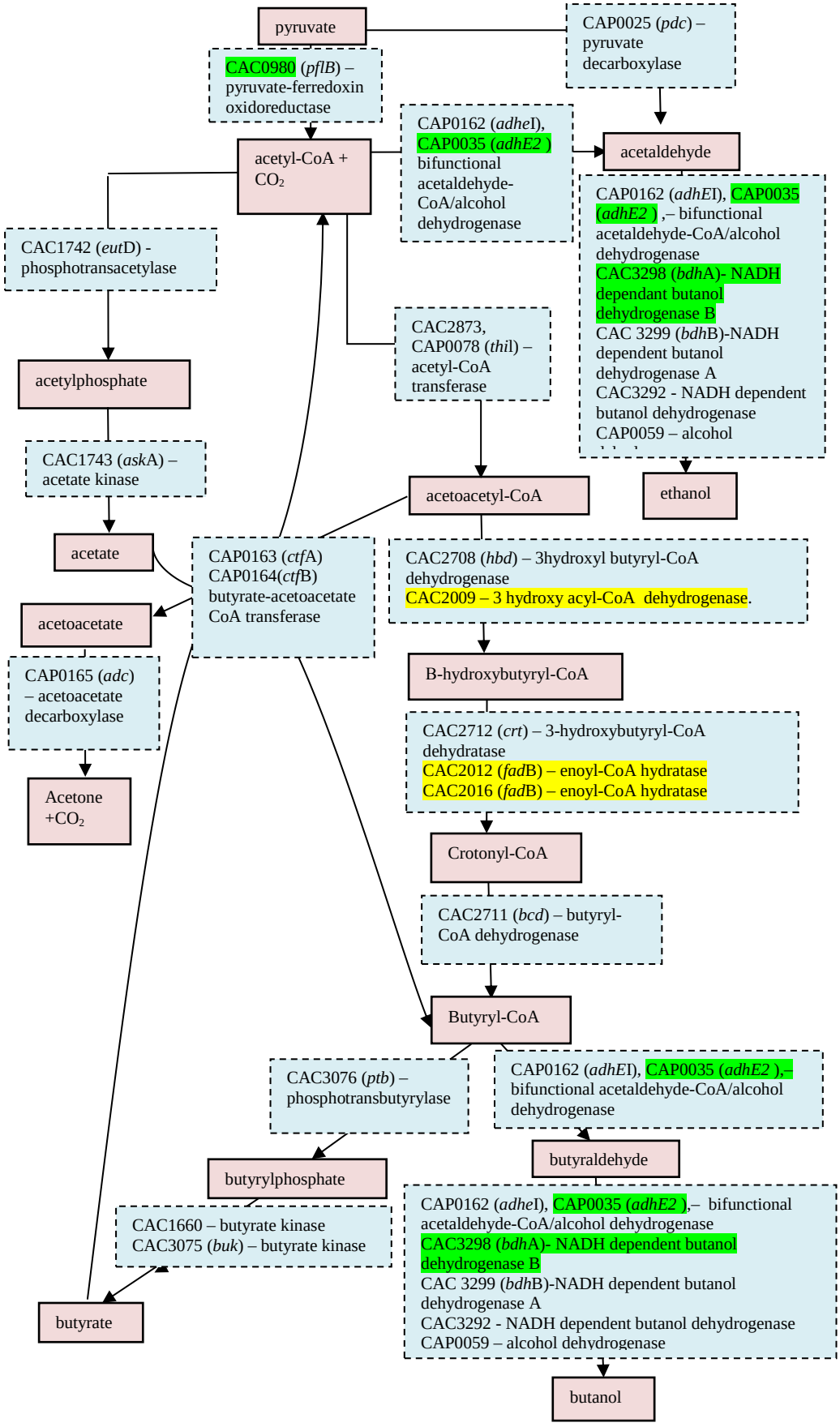


Figure 5.5 Overview of the metabolic pathways involved in acidogenesis and solventogenesis in *C.acetobutylicum* and those genes shown to be differentially regulated in the *agrB* mutant when compared to wild type. The enzymes which are thought to play a role in these pathways are shown in the dashed boxes along with their locus tags (Nolling *et al.,* 2001). Substrates and products are shown in pink boxes. The genes highlighted in yellow were found to be differentially regulated in the wild type strain as compared to the *agrB* mutant at the earlier stage of growth. The green boxes indicate genes which followed the same pattern at the later stage of growth.

5.2.2 Genes involved in sporulation

As the *agrB* mutant was found to be sporulation deficient it is relevant to examine those genes with a direct or indirect involvement in sporulation which have been differentially regulated by *agr*. The sporulation cascade in *C.acetobutylicum* and other bacteria is a complicated pathway with a great many genes involved directly and indirectly. This means that the effect of *agrB* on this cascade could be due to its positive or negative effect on a wide variety of genes

In order to look at some of the genes controlled by the *agr* which may be involved in sporulation it is necessary to look at the whole sporulation pathway. The sporulation pathway in *B.subtilis* has been most extensively studied. Most of the genes involved in the sporulation pathway in this organism are present in the *C.acetobutylicum* genome as described in section 1.8 of the Introduction.

5.2.2.1 Sporulation in *B.subtilis*

The sporulation pathway in *B.subtilis* can be separated into 7 stages as described in figure 1.4. The first stage is the normal vegetative state of the cell followed by stage I the axial filimentation. This stage is characterised by the formation of an axial filament of chromatin along the long axis of the cell.

5.2.2.2 Initiation of Sporulation

In *B.subtilis* the initiation of sporulation is determined by a phosphorelay system which leads to the phosphorylation of Spo0A (see figure 1.4). The first part of this cascade involves the primary kinases KinA and KinB. In response to unknown stimuli they donate their phosphate groups to Spo0F. The phosphotransferase Spo0B then transfers the phosphate from Spo0F to Spo0A.

In *C.acetobutylicum* this phosphorelay system is not present and homologues have not been found in any sequenced clostridia. Spo0A is therefore thought to be activated by an alternative method. Spo0A in *C.botulinum* is phosphorylated by a single orphan kinase CBO1120 (Worner *et al.,* 2006). A number of orphan kinases have recently been investigated as possible activators of Spo0A (Jones *et al.,* 2008). The methodology used was to look at orphan kinases which peaked in expression just before the peak in Spo0A activity. The study used the expression of the *sol* operon (CAP0162-64) as an indicator of Spo0A activity as this is induced by Spo0A~P. The *sol* operon showed an initial 100-fold induction at hour 10 (before *spo0A* reached its maximum expression). This was followed by a second 10-fold induction which was closely followed by the peak in *spo0A* expression. It followed that a candidate kinase would show an increase in expression before the 10 h time point. Of the 4 orphan kinases which showed a peak before the 10 h time point, the proteins encoded by CAC0437 and CAC0903 showed most similarity to CBO1120. The expression of CAC0437 peaked only once before the initial *sol* operon induction and CAC0903 expression peaked before each induction of the *sol* operon (Jones *et al*., 2008). It is interesting to find that of the orphan kinases described it is these two which have been found to be positively regulated by *agr* in this study. CAC0437 was found to be up-regulated 3.4-fold and CAC0903 was found to be up-regulated 25.25-fold in the late wild type strain as compared to the *agrB* mutant at the same stage.

The same study (Jones *et al.,* 2008) also looked at possible two component signal transduction systems which could be candidates as a *spo0A* inducer. Of the 30 annotated histidine kinases which were not orphans, 6 of these were found to peak in expression at 8 h (just before induction of the *sol* operon). Two of these subsequently showed their levels of expression after this time point. Three of the candidates maintained their levels of expression after this time point, one after a dip after 9h. CAC2434 was found to peak at 8 h and then fall back to its initial levels. Its expression then steadily increased again after the second

induction of the *sol* operon. This candidate gene was found to be negatively regulated by *agr* at the early stage of growth in this study as it was found to have 4.2 fold higher expression in the *agrB* mutant as compared to the wild type at the same stage in this study.

It is interesting that *agr* which is involved in a potential global regulatory system would have a positive regulatory effect on other potentially important sensory kinases although it does not appear to have an effect on *spo0A* directly.

The induction of Spo0A formation leads to the initiation of sporulation. Spo0A~P can activate or repress a large number of genes in *B.subtilis* by binding to a 7 bp sequence TGNCGAA where N is any nucleotide. This results in global changes of gene expression affecting around an eighth of the organism's genes. A study comparing a *spo0A* mutant (SKO1) with the wild type strain of *C.acetobutylicum* revealed that 211 genes were differently expressed (Tomas *et al.,* 2003b). This included the down-regulation of solventogenic, sporulation, and carbohydrate metabolism genes and upregulation of flagellar and chemotaxis genes. Spo0A promotes the expression of sigma factors which play the major role in sporulation as described in fig 5.6.

5.2.2.3 Stage I - Axial filament formation

The next stage in sporulation is axial filament formation. This involves the formation of an axial filament of chromatin. Both predivisional chromosomes in the cell elongate along the long axis of the cell. In *B.subtilis* there have been several investigations into the genetic control of this stage of sporulation. Particularly interesting is the coupling of this event and the start of asymmetric division. The DivIVA protein has been studied as possibly being involved at this stage. It is considered a functional homologue of the MinE protein of *E.coli* and restricts the division inhibition proteins in MinCD to the cell poles, ensuring mid-cell division during vegetative growth (Cha and Stewart, 1997; Edwards and Errington,

1997). *B.subtilis divIVA* mutants have been found to be defective in growth and sporulation.

These mutants were often found to have formed anucleate prespores suggesting the point between axial filament formation and asymmetrical division had been disrupted (Thomaides *et al.,* 2001). Incorporation of DNA into the polar prespore compartment is achieved by anchoring one chromosome copy to the distal pole of the prespore compartment. One of the proteins involved in this process is RacA, and DivIVA is responsible for the polar localisation of RacA (Ben-Yehuda *et al*., 2003; Wu and Errington, 2003).

The present study found the *divIVA* gene in *C.acetobutylicum* was up-regulated 2.13-fold in the late wild type strain as compared to the *agrB* mutant at the same stage suggesting that *agr* has a positive regulatory effect on this important gene. The growth of the *agrB* mutant remained normal however. There has been no identification of a *racA* homologue in *C.acetobutylicum* to date. Another gene *soj* (a regulator of *spo0J*) is also implicated in having an effect on axial filament formation, possibly in conjunction with DivIVA and RacA in *B.subtilis* (Wu and Errington, 2003). The N terminus harbors the membrane targeting sequence. The purified protein has been found to stimulate binding of RacA to membranes (Lenarcic *et al.,* 2009).

A homologue of *soj* is present in *C.acetobutylicum* but *agr* had no effect on its expression. It must be noted that *minC* was found to have a 2-fold increase in expression at the later stage of growth in the wild type strain as compared to the *agrB* mutant at the same stage while *minD* was found to have a 2.5 fold increase in expression at the later stage of growth in the *agrB* mutant as compared to the wild type strain at the same stage of growth. MinC is the primary divisional inhibitor of cell division in *E.coli* (de Boer *et al.,* 1991; Hu *et al.,* 1999) and MinD is thought to anchor MinC to the membrane (de Boer *et al.,* 1991). It is interesting therefore to see that *agr* has opposing effects on their expression. The *minC* gene (CAC1248)

in *C.acetobutylicum* was found to be upregulated 2-fold at the later stage of growth in the wild type strain as compared to the *agrB* mutant at the same stage. The *minD* gene (CAC1249) was up-regulated 2 fold at the later stage of growth in the *agrB* mutant as compared to the wild type at the same stage.

C.acetobutylicum has been found not to have an ortholog of the cell division initiation gene *divIC* which suggests some differences in septum formation between *C.acetobutylicum* and *B.subtilis* (Stragier and Losick, 1996).

5.2.2.4 Stage II – Asymmetric division

This stage of sporulation requires a major translocation of the cell's apparatus and is the point when the polar septum is formed. The mother cell and prespore are separated and the compartmentalization of gene expression begins to occur.

During the onset of sporulation Spo0A triggers the formation of Z rings near both poles. This is thought to be mediated by *spoIIE* induction and increased *ftsAZ* expression. There is also evidence that MinC, MinD and SpoVG also play minor roles in selecting the asymmetric division site. Even though the Z rings are formed at both sites along with surface annular structures it is at only one of these sites where the septum is formed (Ben-Yehuda and Losick, 2002; Levin and Losick, 1996; Ryter, 1965). A phenotype associated with *spoII* mutants involves the utilization of both sites to form a cell with two septa. This cell will therefore have two prespores. These mutants can initiate sporulation but do not activate σ^E (Illing and Errington, 1991). Three proteins which are expressed in the mother cell under σ^E control are required to prevent the second asymmetric division (Eichenberger *et al.,* 2001; Pogliano *et al.,* 1999). These 3 corresponding genes are *spoIIM, spoIIP* and *spoIID*. The latter gene has been found to be 2.8-fold up-regulated in the wild type strain at

the later stage of growth as compared to the *agrB* mutant at the same stage of growth indicating that *agrB* has a positive regulatory effect on this gene. This gene is also thought to be required for engulfment of the prespore. The *ftsE* and *ftsW* genes which in *E.coli* are involved in the formation of the divisome with *ftsAZ* were found to be upregulated in the *agrB* mutant at the later stage of growth when compared to the wild type strain at the same stage suggesting that these genes are negatively regulated by *agr*. The *ftsY* gene was found to be upregulated in the wild type strain at the late stage of growth when compared to the *agrB* mutant at the same stage suggesting that *agr* positively regulates this gene.

5.2.2.5 *SpoIIIE*

In *B.subtilis,* during the asymmetric stage of sporulation the cell undergoes division at one of the poles which leaves the origin-proximal one third of a chromosome in the prespore and the other two thirds of the same chromosome and the second chromosome in the mother cell (Wu and Errington, 1998). The remaining origin-distal two thirds of the chromosome therefore, have to be transferred to the prespore. A critical gene involved in this transfer was found to be *spoIIIE* and in *B.subtilis*, mutants in this gene do not transfer the remaining two thirds of the chromosome from the mother cell (Wu and Errington, 1998). SpoIIIE functions as a DNA translocase and localizes to the middle of the sporulation septum. It is capable of forming an effective seal around the DNA during translocation which prevents diffusion of proteins between the prespore and the mother cell (Hilbert *et al.,* 2004; Wu and Errington, 1998). This protein has also been implicated in the stripping off of RNA polymerase, transcription factors, and chromosome remodeling proteins from the DNA during translocation of the chromosome into the forespore compartment (Marquis *et al.,* 2008). SpoIIIE has also been shown to be involved with separation of the chromosomes and engulfment during sporulation (Bogush *et al.,* 2007; Sharp and Pogliano, 1999). CAC0408 described as DNA segregation ATP-ase FtsK/SpoIIIE was found to be up-regulated in the *agrB* mutant 6-fold at the earlier stage of growth when compared with the

wild type strain at the same stage. It was also up-regulated 6.2-fold in the wild type strain at the later stage of growth when compared to the *agrB* mutant at the same stage. This indicates that *agr* down-regulates this gene in the exponential stage of growth and up-regulates the gene in the early stationary phase.

The expression of the locus CAC0039, which is also annotated as a *spoIIIE* was not influenced by the presence of an *agrB* mutation. Another locus, CAC1871, which is also described as an FtsK-like DNA segregation ATPase, followed a similar pattern as it was up-regulated 6.3-fold in the *agrB* mutant at the earlier stage of growth as compared to the wild type strain at the same stage. It was also up-regulated in the wild type strain 204.7-fold at the later stage of growth as compared to the *agrB* mutant at the same stage.

5.2.2.6 Sigma factors

In *B.subtilis* the activation of sigma factors σ^K and σ^E is coupled to asymmetric division. It is at this early stage of sporulation that σ^K, the first sigma factor to be active in the prespore is activated. None of the genes which are implicated to have a role in the activation of σ^K in *B.subtilis* seem to be regulated by *agr*, and *sigF* is not directly regulated. The primary functions of σ^K in *B.subtilis* are to couple prespore and mother cell specific gene expression and to direct the synthesis of σ^G which will be described later in the chapter.

In *B.subtilis* σ^F is involved with the activation of the mother cell specific sigma factor σ^E. In *B.subtilis* σ^E is synthesised in an inactive state and activated by proteolytic cleavage. The gene encoding σ^E is found on the *spoIIG* locus comprising of *spoIIGA* and *spoIIGB* which in *C.acetobutylicum* is CAC1604-05 (Stragier *et al.,* 1984; Trempy and Haldenwang, 1985). The first gene in the operon *spoIIGA* was found to be essential for the processing of σ^E in *B.subtilis* (Jonas *et al.,* 1988; Kenney

and Moran, 1987; Masuda *et al.,* 1990). The second gene *spoIIGB* encodes the σ^E precursor.

This study found no effect of *agrB* on the σ^E precursor (CAC1605) but the SpoIIGA homologue CAC1604 was up-regulated 9.8-fold in the wild type strain at the later stage of growth when compared to the *agrB* mutant at the same stage indicating that this gene is positively regulated by *agr*. In *B.subtilis spoIIGA* is active in response to *spoIIR* whose expression is itself activated by σ^F. *spoIIR* was not found to be significantly differentially regulated in the wild type strain or *agrB* mutant at either stage of growth. If σ^E is activated in the same manner in *C.acetobutylicum* as in *B.subtilis* then this may implicate *agr* as being involved in its activation through the positive regulation of *spoIIGA*. The σ^E regulon has many members in *B.subtilis* including *spoIID* which as mentioned, is required for prevention of a second asymmetric septum and engulfment of the prespore. This gene (CAC2861) was found to be up-regulated 2.8-fold in the wild type mutant at the later stage of growth as compared with the *agrB* mutant at the same stage.

5.2.2.7 Stage III - Engulfment of the prespore

The third stage of sporulation denotes a process by which the prespore is completely engulfed by the mother cell and therefore is completely separated from the mother cell's medium. This stage is initiated by changes in the septum between the prespore and the mother cell. There is a loss of cell wall material initially from the centre of the septal disk and then from the entire septum. The attachments of the septal membrane to the cell wall then migrate to the cell pole which leads to membrane fusion and detachment of the engulfed prespore.

SpoIID has been found to be associated with septal thinning where it degrades peptidoglycan (Abanes-De Mello *et al.,* 2002; Lopez-Diaz *et al.,* 1986). The membrane fusion event at the later stage of engulfment has

also been found to be dependent on the *spoIIIE* locus (Sharp and Pogliano, 1999; Sharp and Pogliano, 2003).

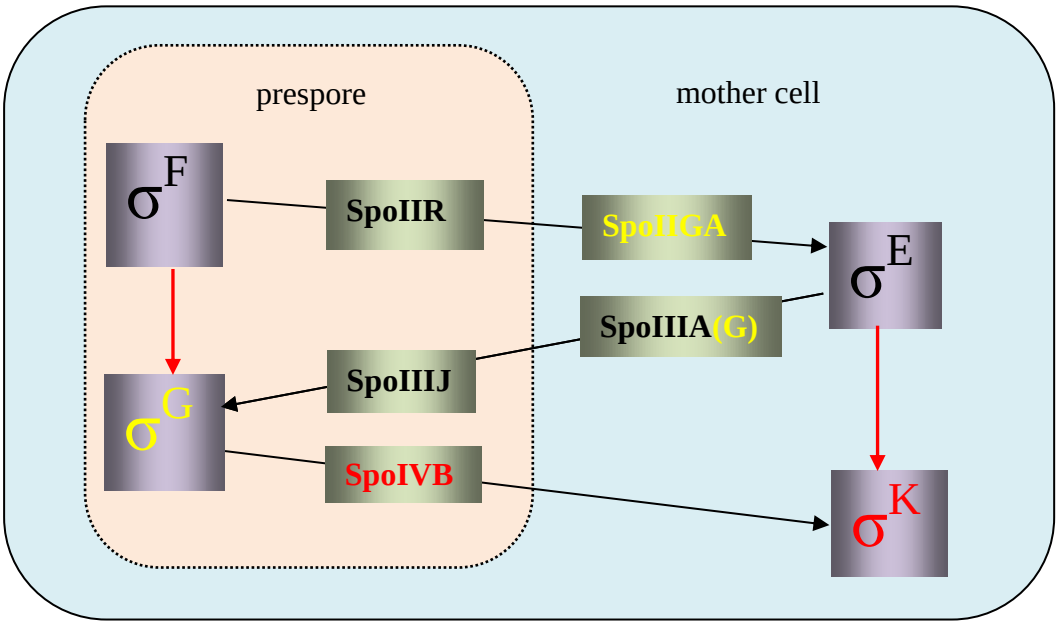


Figure 5.6 The intercompartmental communication between the prespore and mother cell during sporulation in *B.subtilis*, based on a model by Hilbert and Piggot, (2004). All the gene products which exhibited positive expression in the late wild type strain when compared to the late *agrB* mutant are designated printed in yellow. σ^F is the first compartmentalised sigma factor to be expressed during sporulation. This triggers SpoIIR which activates the receptor protease SpoIIGA located in the asymmetric septum. Activated SpoIIGA then processes the inactive precursor pro-σ^E into an active form in the mother cell. σ^E then interacts with RNA polymerase to transcribe the *spoIIIA* operon whose products then interact across the prespore membrane to activate σ^G. σ^G is itself expressed in the prespore under the control of σ^F. The *spoIIIA* operon has been found to have two promoters, one of which controls *spoIIIAG/H* (Guillot and Moran, 2007). The SpoIIIAH protein is thought to act with the SpoIIQ protein which is expressed in the prespore under the control of σ^F. σ^G expression needs SpoIIQ and an as yet unknown signal from the mother cell. SpoIIIJ is also involved in the activation of σ^G although this is expressed vegetatively and is probably found in both compartments. SpoIIIJ has been found to interact with SpoIIIAE (Serrano *et al.,* 2008). The *spoIIIA* operon is likely to maintain σ^G activity by in effect maintaining the metabolic potential between the mother cell and the prespore (Camp and Losick, 2009; Doan *et al.,* 2009). Homologues of those genes illustrated in red have not been found in the *C.acetobutylicum* genome. The search for a potential σ^K is on-going.

5.2.2.8 Sigma Factor G (σ^G)

This late prespore specific sigma factor was found to be up-regulated 8621- fold in the wild type strain of *C.acetobutylicum* at the late stage of growth as compared to the *agrB* mutant at the same stage indicating that the expression of this gene was positively regulated by *agr*. In *B.subtilis* σ^G expression is under the control of σ^F in the prespore. The activation of σ^G coincides with the completion of the engulfment process. The σ^G transcript is detected early along with the σ^E transcript but this transcript is read-through from *sigE* located immediately upstream of *sigG* and is not translated (Chary *et al.,* 2005; Sun *et al.,* 1991). The translation of σ^G in *B.subtilis* takes place when the gene is expressed as a single cistron from a σ^F dependent promoter which is located between *sigE* and *sigG* (Chary *et al.,* 2005, Sun *et al.,* 1991). In *C.acetobutylicum sigE* and *sigG* are adjacent to each other as in *B.subtilis* but no σ^F dependent promoter is present (Paredes *et al.,* 2004). It has therefore been suggested that *sigG* is only expressed as part of the *sigE* operon. This consists of *spoIIGA*, *sigE* and *sigG* which were found to have very similar transcriptional patterns (Jones *et al.,* 2008). This may suggest that they are expressed as a single transcript but a study using Northern blots also found three separate transcripts equating to all three genes, the two genes *spo*IIGA-*sig*E and *sigG* alone (Harris *et al*., 2002). This issue has not been resolved although it is worth noting that the *spoIIGA* gene was found to be up-regulated 9.8 fold in the wild type strain at the later stage of growth as compared to the *agrB* strain at the same stage indicating that it is positively regulated by *agr* along with *sigG*. The expression of *sigE* was not found to be influenced by *agrB*.

The activation of σ^G is described in figure 5.6 and requires the expression of *spoIIIJ* and the *spoIIIA* operon. σ^E directs the expression of the mother cell specific *spoIIIA* operon (Camp and Losick, 2008; Chary *et al.,* 2005; Chary *et al.,* 2006; Chary *et al.,* 2007; Karmazyn-Campelli *et al.,* 2008; Kellner *et al*,, 1996; Partridge and Errington, 1993; Serrano *et al.,* 2003;

Serrano *et al.,* 2004; Serrano *et al.,* 2008; Stragier and Losick, 1996). The *spoIIIA* operon-encoded proteins SpoIIIA to SpoIIIH may form a complex which links the completion of the engulfment process to the activation of σ^G (Broder and Pogliano, 2006; Stragier and Losick, 1996; Camp and Losick, 2008; Doan *et al.,* 2005; Jiang *et al.,* 2005; Kellner *et al*., 1996; Serrano *et al.,* 2008). The product of *spoIIAH* has been found to localise on the outer prespore membrane and interact with the prespore inner membrane protein SpoIIQ. The product of this interaction is a channel through the membrane which is required for σ^G activation (Camp and Losick, 2008; Meisner *et al.,* 2008). Current evidence suggests all the proteins transcribed from the *spoIIIA* operon play a role in forming the SpoIIIA-SpoIIQ complex although their localisation and assembly is unknown. Of all the *spoIIIA* operon genes only *spoIIIAG* was found to be differentially regulated in the wild type/*agrB* comparison. This was up-regulated 1314-fold in the wild type strain at the later stage of growth as compared to the *agrB* mutant at the same stage. The peak in expression of the *spoIIIA* operon has been found to correspond to the largest peak in σ^G activity but an early increase in σ^G activity showed no large induction of *spoIIIA* expression beforehand (Jones et al., 2008).

Although the expression of *spoIIIJ* appears not to be regulated by *agr* there was a 5.5-fold increase in expression of CAC3735 which is described as RNA-binding protein Jag in the wild type strain at the later stage of growth as compared to the *agrB* mutant at the same stage. This is annotated as a *Spo*IIIJ-associated RNA binding protein. This gene is found adjacent to *spoIIIJ* in *C.acetobutylicum*. This gene was not found to be essential for sporulation in *B.subtilis* (Errington *et al.,* 1992).

5.2.2.9 σ^G Regulon

The σ^G regulon in *B.subtilis* consists of genes involved in sporulation, germination and the protection of the spore from damage. One of these genes is *spoVT* which is involved in σ^G dependent regulation (Bagyan *et al.,* 1996). This gene which encodes an *abrB* family transcription factor

has also been implicated in regulating spore coat assembly in *C.beijerinckii* (Shi and Blascheck, 2008). This gene (CAC3649) was upregulated 5.9-fold in the *agrB* mutant at the later stage of growth as compared to the wild type strain at the same stage which indicates that it is negatively regulated by *agr*. This gene was found to be one of two paralogue in *C.acetobutylicum* (Jones *et al*., 2008). It was suggested that this was not the true *spoVT* due to its transcriptional profile while CAC3214 follows a transcriptional pattern more appropriate for this gene. CAC3214 was annotated as a regulator of stationary/sporulation gene expression in the same family as *spoVT*. It was not found to be significantly upregulated in a differential manner at either stage of growth.

Many other genes which are regulated by σ^G in *B.subtilis* do not seem to be regulated by *agr*. σ^G has been found to regulate itself in an autocatalytic loop *spoIVB* which is involved in σ^K activation and *spoVA* which is required for the uptake of dipicolonic acid into the prespore from the mother cell (Errington and Mandelstam; 1984, Moldover *et al.,* 1991; Tovar-Rojo *et al.,* 2002). It has also been found to be involved in the regulation of *ssp* genes which encode small acid-soluble proteins (SASPs). These are the predominant proteins in the spore core and protect DNA from adverse conditions which the spore could encounter. They are also thought to provide a source of amino acids upon germination. The only differentially expressed gene thought to encode an SASP was CAC1620 which was up-regulated 4.3-fold in the *agrB* mutant at the late stage of growth as compared to the wild type strain at the same stage. It is, therefore, thought to be negatively regulated by *agr*. Finally CAC3081 annotated as *sleB* was found to be up-regulated 678.5-fold in the wild type strain at the later stage of growth as compared to the *agrB* mutant at the same stage. This gene which encodes a spore cortex lytic enzyme vital for germination has been found to be under the control of σ^G in *B.subtilis* (Moriyama *et al.,* 1999).

5.2.2.9 Sigma factor K (σ^K)

Following activation of σ^G in the prespore σ^K becomes active in the mother cell. The genes in the σ^K regulon are mainly involved in the spore coat formation and spore maturation. The expression of the gene annotated as *sigK* in *C.acetobutylicum* (CAC1689) showed no differential expression in the wild type or *agrB* mutant at either stage of growth. Furthermore this gene has not been found to follow the expected expression profile of *sig*K (Jones *et al.,* 2008). The same study found that changes to the level of expression of *spoIVFB* (CAC1253) the putative σ^K processing enzyme did not exceed the cut off criteria either.

Further investigation into the expression of genes in the σ^K regulon which included *yabG* and *spsF* found they were expressed in mid- and late-stationary phase, respectively, and did not confirm the existence of a functional *sigK* in *C.acetobutylicum* (Jones *et al.,* 2008). Interestingly *spsF* was found to be up-regulated 2.3-fold in the wild type strain at the early stage of growth when compared to the *agrB* mutant at the same stage indicating positive regulation by *agrB.* This pattern was not repeated at the later stage of growth.

Other orthologues of the *B.subtilis* σ^K regulon have been found to be expressed in what could be an *agrB*-dependent manner. The gene which is annotated as *spoVD* (CAP0150) is found on the pSOL-1 plasmid and was up-regulated 953.9-fold in the wild type strain at the late stage of growth as compared to the *agrB* mutant at the same stage. This indicates that it may be indirectly positively regulated by *agr* in the late stage of growth. This gene is required for spore maturation in *B.subtilis* (Daniel *et al.,* 1994; Fan *et al*., 1992). A *cotS* gene has also been found to be up-regulated 3106 fold in the wild type strain at the late stage of growth as compared to the *agrB* mutant at the same stage. This gene was found to be regulated in a σ^K dependent manner but a mutant of *cot*S does not produce a sporulation deficient phenotype (Abe *et al.,* 1995; Takamatsu *et al.,* 1998). Another *cot* gene described as *cotF* (CAC2683) was up-

regulated 3.8 fold in the *agrB* mutant at the earlier stage of growth as compared to the wild type strain at the same stage in this study. *cotF* is also thought to be controlled in a σ^K dependent manner (Takamatsu *et al*., 1998).

Although the comparison with known *B.subtilis* genes which are involved in sporulation is a useful tool for identifying genes which *agr* may regulate there are distinct differences in the sporulation pathways of both organisms. The number and diversity of detectable sporulation genes in *C.acetobutylicum* is much smaller than with *B.subtilis*. The distinct family of histidine kinases found in *B.subtilis* which regulate the initiation of sporulation (KinA/B and RAP) are missing in *C.acetobutylicum*. The major genes involved in the phosphorelay system of stage 0 in *B.subtilis* (*spo0B* and *spo0F*) are also missing in *C.acetobutylicum*. In the later stages of sporulation *spoVF*, *spoVK* and *spoVM* are not present in *C.acetobutylicum* (Jones *et al.,* 2008).

Analysis of the putative *agr* regulon has revealed several important genes which have been implicated in sporulation in *B.subtilis*. Many of these genes are implicated in the early and mid-stages of the sporulation pathway. This provides evidence as to why sporulation is deficient in the *agrB* mutant as found in the chapter 4.

5.2.2.10 Uncharacterised sigma factors

A number of sigma factors have been investigated previously for a possible role in sporulation (Jones *et al.,* 2008). CAC1770, described as a putative RNA polymerase sigma factor *sigI* was found to be up-regulated 1334 fold in the *agrB* mutant at the early stage of growth as compared to the wild type strain at the same stage. CAC1509 is described as a specialised polymerase subunit and was found to be up-regulated 12.4 fold in the *agrB* mutant at the early stage of growth as compared to the wild type strain at the same stage. The expression of these potential

sigma factors has been found to be below the designated cut off value to be implicated in sporulation (Jones *et al.,* 2008).

CAC0550 has been found to be expressed at its highest levels in the early exponential phase of growth (Jones *et al.,* 2008). It was at this stage that a 15 fold difference of expression was found in the early wild type strain as compared to the *agrB* mutant at the same stage. Although this may not be involved in sporulation as a result it is nevertheless an interesting potential sigma factor for future studies.

CAP0167 was found to be up-regulated 5.9 fold in the wild type strain at the early stage of growth as compared to the *agrB* mutant at the same stage. This gene was found to have a similar expression profile to *spo0A* (Jones *et al.,* 2008). The gene has also been predicted to be part of a bicistronic operon along with CAC0166 driven from a promoter of the σ^F/σ^G family (Paredes *et al.,* 2004). Antisense RNA was used to target this gene and CAP0166 and the asCAP0167 strain was found to have altered cigar shape morphology after 72 h of growth (Jones *et al.,* 2008). As this study found CAP0167 showed a greater identity to other sigma factors than *sigE*, *sigF* and *sigG* it was suggested that this was a novel sporulation related σ factor in clostridia. CAP0166 was not found to be differentially regulated in the wild type or *agrB* mutant at any stage of growth.

5.2.2.11. pSOL-1 genes

The genes described in this section, (designated CAP genes) belong to the pSOL-1 megaplasmid which is known to be essential for the production of solvents and normal sporulation. Interestingly at the early stage of growth there were 13 genes upregulated in the *agrB* mutant as compared to the wild type strain at the same stage and 4 genes were found to be up-regulated in the wild type as compared to the *agrB* mutant at the same stage.

The opposite pattern occurred at the later stage of growth with 16 genes up-regulated in the wild type strain as compared to the *agrB* mutant at the same stage and 5 genes up-regulated in the *agrB* mutant as compared to the wild type strain at the same stage.

At the earlier stage of growth most of the genes up-regulated in the early stage *agrB* mutant as compared to the wild type strain encode hypothetical proteins including CAP0137, CAP0138, CAP0158 and CAP0159.

The highest level of up regulation as compared to the wild type strain at the same stage was with CAP 0137, encoding an enzyme similar to UDP-glucuronosyltransferases. This was up-regulated 1230-fold. CAP0146 was up-regulated 716 fold and was described as a *hypF* homologue, a gene involved in the production of hydrogenases (Rangarajan et al., 2008). CAP0048, described as encoding a methyl accepting chemotaxis protein (see fig 5.7) was up-regulated 2.4-fold. Two other possibly linked genes are CAP0146 and CAP0147 annotated as encoding a YdfR family protein and a hypothetical protein, which were up-regulated 715- and 688-fold respectively.

At the early stage of growth the 4 genes up-regulated in the wild type strain as compared to the *agrB* mutant at the same stage included CAP0058, annotated as a RplA related protein, CAP0151, described as encoding an lysozyme and CAP0167, encoding a *sigF*/*sigE* family sigma factor. As previously stated the down regulation of this gene by antisense RNA has been found to affect sporulation (Jones et al., 2008).

At the later stage of growth the 5 genes up-regulated in the *agrB* mutant as compared to the wild type strain at the same stage included CAP0035, the second *adhE* gene found in *C.acetobutylicum* which is thought to be involved in solvent production under certain conditions as discussed previously. Two transcriptional regulators, CAP0107 and CAP0127 were up-regulated 3- and 4-fold respectively.

The genes up-regulated in the wild type strain at the later stage of growth as compared to the *agrB* mutant at the same stage also mainly encoded hypothetical proteins but perhaps the most important gene to be up-regulated is *spoVD*. In *B.subtilis* this gene is expressed in the mother cell compartment during sporulation and plays a role in cortex synthesis (Vasudevan et al., 2007). In *B.subtilis* this gene is under the control of *spoIIID* which has been suggested to be under σ^E control in *C.acetobutylicum* (Jones et al, 2008). There was however no significant up-regulation of either of these genes in the wild type strain at either stage of growth.

Four genes, CAP0109, CAP0130, CAP0137 and CAP0138 were found to be up-regulated at the early stage of growth in the *agrB* mutant as compared to the wild type strain. At the later stage of growth they were found to be up-regulated in the wild type strain as compared to the *agrB* mutant at the same stage of growth. As found with other genes this indicated that the *agr* could be regulating these genes indirectly.

5.2.3 Cell motility, chemotaxis and signal transduction mechanisms

5.2.3.1 COG class N – Cell motility and secretion

This group of genes involved with motility and cell secretion was found to be largely down regulated at the transitional phase as the cells start to reach solventogenic growth. Some of the genes were then found to be up-regulated again in the late stationary phase (Alsaker and Papoutsakis. 2005). Motility is associated with greater solvent production due to chemotaxis towards required substrates (Gutierrez and Maddox. 1987).

At the earlier stage of growth one gene was up-regulated significantly in the *agrB* mutant as compared to the wild type strain. This was CAC1389, encoding a ChW repeat containing protein. This type of protein belongs to

a novel protein family which is specific to *C.acetobutylicum*. The proteins contain a conserved tryptophan but their function has not been characterised and may be unique to the organism's physiology (Sullivan *et al.,* 2007). These types of proteins were also found to be up-regulated in the early wild type as compared to the *agrB* mutant at the same stage (CAC3279 and CAC3273).

Of the 9 genes up-regulated in the wild type at the earlier stage of growth as compared to the *agrB* mutant at the same stage, 6 were described as cell adhesion domain containing proteins, 5 of which had the highest fold difference in up-regulation. CAC1634 is annotated as a flagellin and was up-regulated 3.4 fold. There were also two ChW repeat containing proteins which followed this pattern.

In the later stage of growth the *agrB* mutant had 4 genes up-regulated compared to the wild type strain at the same stage and the wild type strain had 2 genes up-regulated as compared to the *agrB* mutant at the same stage. The highest level of up-regulation in the *agrB* mutant at the later stage of growth compared to the wild type strain at the same stage was with CAC0304, encoding a chemotaxis motility protein A (*motA*). The three other genes following this pattern were CAC2165 (*flgB*), CAC2164 (*flgC*) and CAC2146 (*flfF*). Both *flgB* and *flgC* encode flagellar proximal rod proteins (Zuberi *et al.,* 1991). CAC2146 (*flhF*) is annotated as a flagellar biosynthesis regulator and in *B.subtilis* is located on the *che*/*fla* operon (Kirsch *et al.,* 1994). The proteomic and biochemical analysis of *flh*F in *B.subtilis* strain 168 demonstrated that this gene is dispensable for protein secretion and plays a minor role in cell motility (Zanen *et al.,* 2004).

The 2 up-regulated genes in the wild type as compared to the *agrB* mutant at the same stage, were CAC3274 a ChW repeat containing protein which was also up-regulated at the earlier stage of growth and CAC3565 encoding a cell adhesion domain containing protein.

This data show that *agr* is likely to be involved in the positive regulation of cell adhesion proteins and the negative regulation of flagellar proteins. Figure 5.7 shows where some of the differentially regulated gene products are involved in the motility apparatus.

5.2.3.2 COG Group T – Signal transduction mechanisms

As *agrB* is a putative signal processor for a global regulatory quorum sensing system, one might expect its disruption to have an effect on other signalling transduction mechanisms.

When compared to other COG groups this group has a large number of genes which are significantly up-regulated in both the *agrB* mutant and the wild type. At the early stage of growth there were 12 genes up-regulated in the *agrB* mutant compared to the wild type strain at the same stage and 17 genes up-regulated in the wild type strain as compared to the *agrB* mutant at the same stage. Two of the highest levels of up-regulation in the *agrB* mutant compared to the wild type strain were with CAC2763 and CAC3397 which were both described as encoding methyl accepting chemotaxis proteins and were up-regulated by around 1200 fold. CAC1507, described as a sensory transduction histidine kinase (*phoR*) was up-regulated 1102 fold. This gene is involved in the regulation of genes belonging to the *pho* regulon which is involved in overcoming phosphate deficiency in *B.subtilis* (Eldakak and Hulett, 2007). CAC2958, described as a PTS system enzyme IIBC component was up-regulated 762-fold. Most of the other genes found to be up-regulated in the *agrB* mutant compared to the wild type strain at the same stage were histidine kinases.

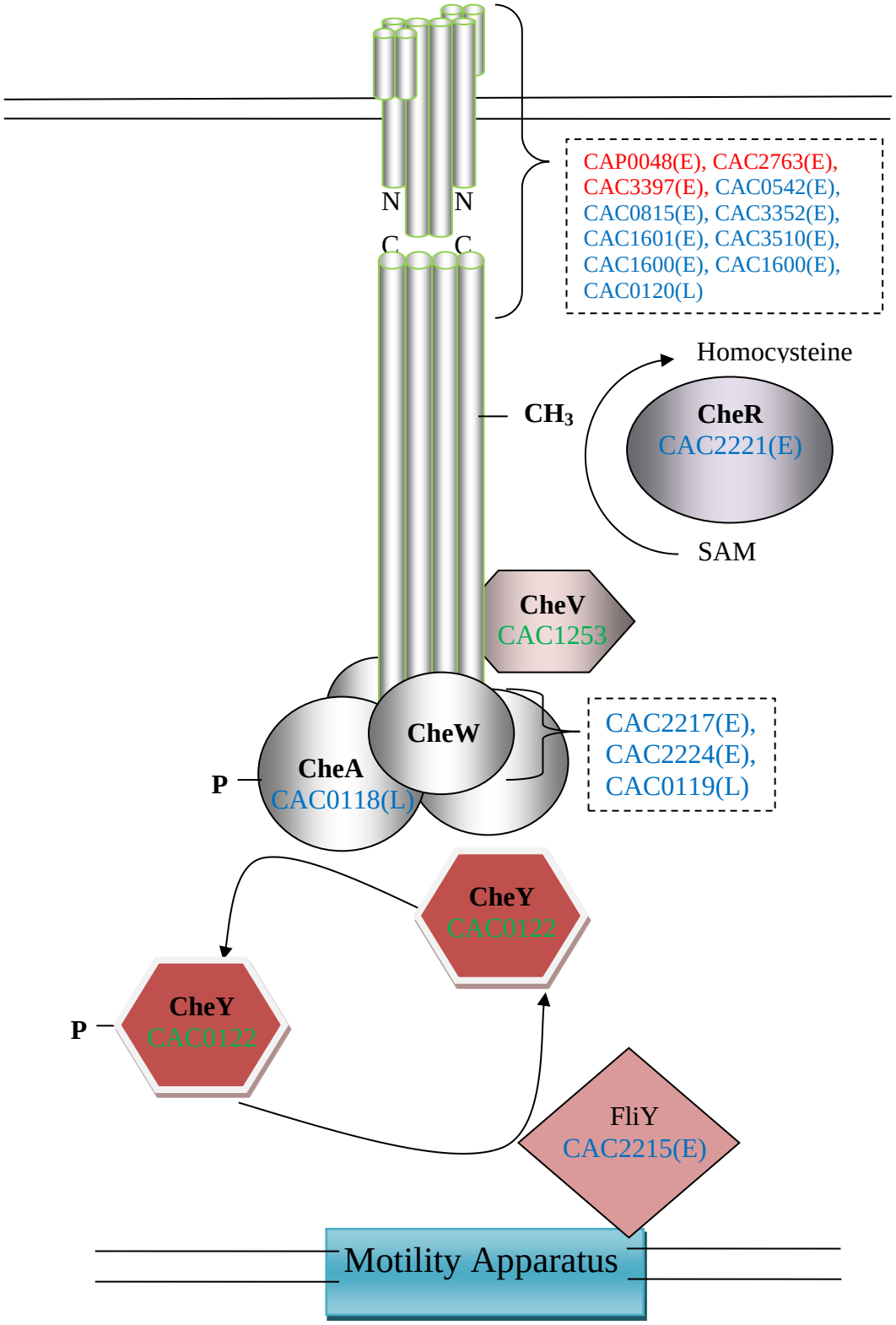


Figure 5.7 The chemotaxis apparatus and the locus numbers of orthologues which were found to have different expression in the *agrB* mutant when compared to the wild type strain**.** Text highlighted in red indicates orthologues for which *agr* appears to have a negative effect on their expression. Blue text indicates those orthologues for which *agrB* has a positive influence on their expression. Those orthologues where *agr* had a positive effect on their expression at the early stage of growth but a negative effect on expression at the later stage of growth are shown in green. The letter E (early) or L (late) designates the stage in growth where the positive or negative effect on expression occurred.

In the early stage of growth, the wild type strain had many genes involved in chemotaxis up-regulated as compared to the *agrB* mutant at the same stage. The two genes with the highest fold up-regulation were two methyl accepting chemotaxis proteins CAC0542 (7.6-fold) and CAC0815 (6.6-fold). In all there are 6 examples of genes encoding this type of methyl accepting chemotaxis protein following the same pattern in the wild type strain. CAC0080, encoding the putative histidine kinase in the *agr* system is not surprisingly up-regulated 4.1 fold in the early wild type strain as compared to the *agrB* mutant at the same stage suggesting that it is positively regulated by the *agrB* gene. CAC2209 was up-regulated 2.35-fold and is annotated as a carbon storage regulator (*csr*A). In *E.coli* this gene down-regulates CstA, a protein involved with nutrient scavenging during carbon limitation (Dubey *et al.,* 2003). However this study found that CstA was also up-regulated in the wild type strain at the early stage of growth as compared to the *agrB* mutant at the same stage.

There were 5 other chemotaxis genes (CAC1233, CAC217, CAC2224, CAC2221 and CAC0112) up-regulated between 2.1 and 2.6 fold in the wild type strain at the early stage of growth as compared to the *agrB* mutant at the same stage. These included the gene encoding the main chemotaxis response regulator CheY which is responsible for controlling flagellar rotation in *B.subtilis* (Welch *et al.,* 1994). Interestingly the same gene, along with CAC1233 (*cheV*) is also up-regulated 4.8- and 5.6-fold respectively in the *agrB* mutant at the later stage of growth compared to the wild type strain at the same stage. This could be due to the global regulation of *agr* and its effects on other regulatory systems in the cell.

Previous analysis indicates a bimodal expression of motility genes whereby they are up-regulated during exponential stage and the late stationary phase (Jones *et al*., 2008). It is hard to provide further evidence without quantitative analysis but it seems with *cheY* at least that *agr* has a positive effect during exponential growth and a negative effect during the transitional phase of growth.

At the later stage of growth the gene CAC0118 which is annotated as CheA is up-regulated nearly 5000-fold in the wild type strain as compared to the *agrB* mutant at the same stage. In *B.subtilis* chemotaxis receptors control the phosphorylation of CheA which itself acts as a substrate for CheY (Szurmant and Ordal, 2004). Why this is positively regulated at the same time as *cheY* is negatively regulated raises an interesting question. It is not surprising that another gene; *che*W (CAC0119) is up-regulated 2182-fold along with *che*A as compared to the *agrB* mutant at the same stage, as this class of protein is thought to couple CheA to receptors (Boukhvalova *et al.,* 2002).

Many of these proteins involved with chemotaxis are shown in Figure 5.7. It is not surprising that a putative global regulator should be involved in the regulation of chemotaxis genes.

5.2.4 Carbohydrate transport and metabolism

The genes associated with this group were largely up-regulated in the *agrB* mutant at the early stage of growth as compared to the wild type strain at the same stage and in the wild type strain at the late stage of growth as compared to the *agrB* mutant at the same stage.

At the early stage of growth there were 14 genes up-regulated in the *agrB* mutant as compared to the wild type strain at the same stage and 4 genes up-regulated in the wild type strain as compared to the *agrB* mutant.
CAC0664 had the highest fold up regulation in the *agrB* mutant as compared to the wild type strain at the early stage of growth. This gene which is annotated as encoding a sugar binding periplasmic protein was up-regulated 1058-fold. There were 7 other genes up-regulated over 200-fold in the *agrB* mutant including L-arabinose isomarase (*araA*).

The highest level of up regulation in the wild type strain as compared to the *agrB* mutant at the earlier stage of growth was for CAC2807 described as encoding an endo-1,3(4)-beta-glucanase family protein 16 enzyme which was up-regulated 6.7-fold. This was followed by CAC0712,

encoding a phosphoglucomutase (*pgm*) and genes encoding 2 nucleoside-diphosphate sugar epimerases.

At the late stage there were 6 genes up-regulated in the *agrB* mutant as compared to the wild type strain at the same stage and 12 genes up-regulated in the wild type strain as compared to the *agrB* mutant at the same stage.

At the later stages of growth the *agrB* mutant had a higher level of expression as compared to the wild type strain for CAC1354, a phosphotransferase system IIA component (6.33-fold). Other up-regulated genes following the same pattern included those encoding a phospho glycerate mutase, a methylglyoxal synthase (*mgsA*), 2 sugar metabolism transcriptional regulators and a transketolase.

In the wild type strain at the later stage of growth CAC0664 was up-regulated 10130-fold as compared to the *agrB* mutant. This showed that *agrB* may be acting in an indirect manner to regulate this gene.

Another interesting gene up-regulated at this stage of growth in the wild type strain as compared to the *agrB* mutant was CAC2239 described as a glycogen synthase gene (*glgA*). This gene was up-regulated 9910-fold. A gene associated with granulose formation, *glg*A has been found to be up-regulated during the transitional phase of growth (Jones *et al*., 2008). Granulose builds up in *C.acetobutylicum* during the stationary phase of growth as a storage product.

Other genes following the same pattern were two putatively linked genes CAC0660 and CAC0665 which are described as encoding permease genes and were up-regulated 8460- and 3444-fold respectively as compared to the *agrB* mutant at the same stage. Two other permease genes, CAC0770 and CAC3670 (*amy*C) were also up-regulated 2494- and 2001-fold respectively as compared to the *agrB* mutant at the same stage. CAC1349, encoding an aldose-1-epimerase (*gal*M) and CAC6358,

encoding a xylanase/chitin deacetylase were up-regulated 2225- and 1241-fold as compared to the *agrB* mutant at the same stage. These results suggest that *agr* has a largely negative effect on sugar metabolism genes at the late stage of growth.

5.3 Discussion

This study has shown that *agr* appears to be implicated in both the positive and negative regulation of a wide variety of genes in *C.acetobutylicum*. The question as to whether this is a true quorum sensing system has still to be answered.

The qualitative data obtained using this microarray technology has to be verified by quantitative data which could be obtained using QRT-PCR. Due to time constraints this analysis was not performed. The results do however, give an idea of which genes may be under the control of this putative *agr* system.

5.3.1 The effect of the *agrB* deletion on sporulation

The deletion of the *agr* homologues of *C.acetobutylicum* has been found to disrupt sporulation in the organism. There were no spores found in the plate assays with the *agrC* and *agrA* mutants and a significantly reduced spore count when the *agrB* gene was disrupted. There was also disruption of sporulation in all the complemented strains which suggests the mutant could not be complemented with these vectors. When the empty vector used for complementation was transformed into the cell however it also affected sporulation to a large degree suggesting it was the presence of the vector itself which caused this effect. pIMP1 has been found to affect solvent production in *C.acetobutylicum* (Walter *et al.,* 1994).

The most probable cause is the metabolic burden the plasmid puts on the cell. No equivalent study has been reported looking into the effects of plasmids on sporulation in this species.

The mutants were screened in order to make sure there were no plasmids present using PCR. Since until recently there have been very few mutations obtained in *C.acetobutylicum,* it is unclear how easy it is to complement a mutation in this species. The non-solventogenic mutant strain M5 has been complemented (Lee *et al*., 2009).

The microarray provided additional information as to why knocking out these *agr* genes had such an effect on sporulation. This is assuming that *agrB* is part of a regulatory system which incorporates *agrC* and *agrA*. Knocking out the *agrB* gene was found to affect the expression of a wide variety of genes which have been implicated in the regulation of sporulation. Examples of these genes are listed in table 5.2.

The regulation and transcriptional programme of sporulation in *C.acetobutylicum* is still largely unknown and much research is needed in order to fully understand the sporulation process in the organism. This study has found evidence to suggest that *agr* is involved in the regulation of sporulation in *C.acetobutylicum*.

Strain and time of differential expression	Fold increase in expression	Locus Number	Gene description
agrB mutant (Early stage of growth)	3.5	CAC0437	Sensory transduction histidine kinase
	4.2	CAC2434	histidine kinase
	5.3	CAC2909	Spore coat protein *cotS* related (diverged)
	23.5	CAC2910	Spore coat protein cotS related
Wild type (Early stage of growth)	2.3	CAC2187	Sialic acid synthase *spsE*
	2.3	CAC2190	Spore coat polysaccharide biosynthesis protein F *spsF*
	5.0	CAC1335	Spore coat protein F (*cotF*) related protein
agrB mutant (Late stage of growth)	25.3	CAC0903	Sensory transduction histidine kinase
	2.2	CAC1251	Cell cycle protein *ftsW* (*rodA*)
	2.6	CAC0497	Cell division ATP binding protein *ftsE*
	2.5	CAC1249	Septum site-determining protein *minD*, ATPase
	5.9	CAC3649	Possible stage V sporulation protein T, transcriptional regulator *abrB* homologue (*spoVT*)
Wild type (Late stage of growth)	2.0	CAC1248	Septum formation inhibitor *minC*
	2.8	CAC2861	Sporulation protein *spoIID*
	3106.0	CAC2906	Spore coat protein *cotS* related
	9.8	CAC1694	Sigma factor E reprocessing enzyme, *spoIIGA*
	8621.0	CAC1696	Sporulation sigma factor *sigG*
	1314.0	CAC2087	Stage III sporulation protein AG, *spoIIAG*
	5.5	CAC3735	RNA binding protein Jag, *spoIIIJ*-associated
agrB mutant (Early stage of growth) + wild type (Late stage of growth)	6.0 + 6.2	CAC0408	DNA segregation ATPase FtsK/SpoIIIE
	6.3 + 204.7	CAC1871	FtsK-like DNA segregation ATPase
agrB mutant (early stage of growth) + *agrB* mutant (Late stage of growth)	3.7 + 4.3	CAC1620	Small acid soluble spore protein

Table 5.2 Examples of genes involved in sporulation whose expression is affected by the mutation of *agrB*

There were some genes which have been found to be expressed much later in the growth phase which appear to be positively regulated by *agr* at the transitional stage of growth such as *sigG*. Jones *et al.* (2008) found this gene to be up-regulated at +40 hours into growth. Why this gene was found to be differentially regulated at this stage of growth should be investigated further.

The growth of the organism in this study was conducted in a simple batch culture which was not buffered. This may have caused the transitional phase to occur at an earlier period of time which could lead to the expression of *sigG* at an earlier stage of growth. It could also be due to the

fact that there was a large scale difference in very low expression of this gene at the early phase of growth. This could be confirmed with quantitative analysis using RT-PCR. The same could be said for the stage III gene *spoIIIAG*.

It is tempting to speculate by which mechanism the *agr* system has an effect on sporulation. Spo0A is the known master regulator of sporulation and solventogenesis so the putative *agr* system may affect the functioning of this protein. Knockout of *agrB* had no effect on the expression of *spo0A* but the *agr* system may have an effect on its phosphorylation, a prerequisite for Spo0A function.

As previously described the expression profiles of potential histidine kinases which could be implicated in the phosphorylation of Spo0A have been examined (Jones et al, 2008). If the *agr* system had a direct effect on these kinases then it should indirectly have an effect on the level of phosphorylation of Spo0A. If this was the case then one would expect to see down regulation of genes which are positively regulated by Spo0A~P. This is certainly the case with some of these genes as shown in figure 5.6. As a result this study shows the burgeoning evidence that the *agr* system may play a major role as a regulator of sporulation and possibly solventogenesis in *C.acetobutylicum*. Furthermore it provides some evidence that sporulation, and possibly solventogenesis could be regulated in a cell density dependent manner in this species.

5.3.2 The effect of the *agr* deletion on solventogenesis

Solventogenesis in *C.acetobutylicum* is perhaps the most extensively studied aspect of this organism's physiology. This is mainly due to the commercial benefits of solvent production. The level of product formation in the wild type strain was not found to be consistent with published levels. This is most likely to be due to the medium the cells were grown in. This medium was based on a recipe used to achieve consistent levels of growth by the organism in batch culture and does not seem to be the most

suitable medium for product formation. It is, therefore, not possible to suggest that the small differences seen in the product formation, particularly with the *agrA* mutant, were significant. It is tempting to suggest that disrupting the *agr* homologues in *C.acetobutylicum* may have no effect on product formation.

The complementation vectors also had little effect on solventogenesis when introduced into the mutant and wild type. This finding would be expected if the *agr* genes had little effect on product formation although the complementation vectors appear to have not worked effectively, as described below.

The data obtained from the microarray suggest that most of the significant genes found to be involved with normal product formation were not influenced by *agr*. There were, however, some examples of genes differentially expressed which have been found to be involved in metabolic pathways. Regulation was found to be affected by the knockout of *agrB* as can be seen in figure 5.5. CAC0980 was found to be up-regulated 3-fold at the later stage of growth in the *agrB* mutant as compared to the wild type strain at the same stage of growth. The product of this gene is involved in the conversion of pyruvate into acetyl-CoA. CAP0035, described as encoding a bifunctional acetaldehyde-CoA/alcohol dehydrogenase (*adhE*) is possibly involved in the formation of both ethanol and butanol (Nolling *et al.,* 2001). This gene was found to be up-regulated at the later stage of growth in the *agrB* mutant as compared to the wild type strain at the same stage of growth suggesting that at the later stage of growth it is negatively regulated by *agr*.

There has been some interest in *adhE*2 as it was found to be responsible for butanol production during alcohologenic cultures of *C.acetobutylicum*. Fontaine *et al.* (2002) found *adhE2* was specifically expressed during this type of growth by Northern blot although Jones *et al.* (2008) found this gene to be part of a cluster of genes up-regulated during the transitional phase of growth under normal growth conditions. This is the second *adhE*

gene to be discovered in *C.acetobutylicum* and is not, therefore, the only gene involved at this stage of product formation. A second gene CAC3298, encoding a NADH dependent butanol dehydrogenase B, was up-regulated 2-fold in the late wild type strain and may be one of the 6 genes along with CAP0035 involved in the conversion of acetaldehyde and butyraldehyde to ethanol and butanol, the results suggest CAP0035 is negatively regulated by *agr*.

There were 3 other genomic genes which were found to be up-regulated in the early wild type strain indicating they are positively regulated by *agrB*. CAC2009, described as encoding a 3-hydroxy acyl-CoA dehydrogenase, CAC2012 and CAC2016, both described as encoding enoyl-CoA hydratases (both annotated as *fadB*) and thought to be involved in the conversion pathway from acetoacetyl-CoA to butyryl-CoA (Nolling *et al.,* 2001). These genes in the COG classification are classified as being involved in lipid metabolism (functional category I). Furthermore these genes are part of a locus which has been found to be up-regulated in stationary phase (Jones *et al.,* 2008).

This study found the same locus was positively regulated by *agr* at an earlier stage of growth but not at a later stage. This will be discussed later in this chapter. The *agrB* mutation was found to have no effect on the regulation of the vast majority of the genes involved in product formation. This includes the genes present in the *sol* operon (*adhE ctfB* and *ctfA*). The fact that so many genes involved in product formation do not appear to be under the regulation of *agrB* suggests that the putative *agr* system in *C.acetobutylicum* does not influence product formation.

A full analysis of product formation using the correct medium in a buffered bioreactor would have to be performed in order to confirm this hypothesis. Ideally this would be performed using previously published methods over the full course of the organism's growth so the data could be directly compared with previously published information.

5.3.3 The effect of the *agr* knockout on motility genes

The mutation of *agrB* was found to affect the expression of a variety of genes involved in motility. This is not surprising since the overexpression of Spo0A has a negative effect on motility (Alsaker *et al.,* 2004). The chemotaxis and motility genes have been found to be clustered together in a bimodal expression pattern where there was increased expression during both exponential and late stationary phase of growth (Jones et al., 2008). This indicates that these genes may be regulated in a cell density dependent manner. This finding gives justification to look at the use of motility assays to identify if mutation of *agr* genes has a phenotypic effect on motility.

5.3.4 Does agr act in a cell density dependent manner*?*

The *agr* system has been found in many different firmicutes and it has been suggested that the regulatory network controlled by the *agr* system is flexible and not analogous in all species (Wuster and Babu, 2008). This publication compared genes which were differentially expressed in knockout mutants of *L.plantarum*, *E.faecalis*, *S.aureus* and *S.epidermidis*. They defined clusters of orthologous genes and looked to see if any were differently expressed in all the mutants studied. They found no clusters differentially expressed. It therefore appears that there is no propensity for orthologous genes to be regulated by the *agr* system in the selected species (Wuster and Babu, 2008). This study also found RNAIII only occurred in staphylococcal species suggesting it is an evolutionary unique mechanism of regulation for this species. It is perhaps not surprising that because the *agr* system is so wide spread that is capable of regulating a wide variety of genes using other downstream regulators.

Some the genes which appear to be regulated by the putative *agr* system have been mentioned in the previous sections. The *agr* system is a cell density dependent regulatory system, so it would make sense that some genes would be up-regulated or down regulated according to the point in the growth phase. In *S.aureus,* for example, cell surface adhesins are up-

regulated at low cell densities and down-regulated at high cell densities in an *agr* dependent manner.

Ideally it should be possible to measure the differences in gene expression over several different time points in the growth phase but it may still be possible to see a difference in a gene expression between two time points. In this study it was possible to show which genes were potentially down regulated and up-regulated by the presence or absence of a functional *agrB* gene, but to measure whether the regulation was in a cell density dependent manner it is necessary to look at genes which increase in positive or negative regulation over the growth phase. For example, in the case of CAC0078, the *agrB* gene itself increases from 72.4-fold to 1272-fold when compared to wild type at the two stages of the growth phase analysed.

There were 11 genes which followed this pattern which, although positively regulated by the presence of the *agrB* gene, had a higher differential level of expression in the later stage of growth than the early stage. Of those genes which were found to be down-regulated by *agr*10 had an increase of this regulation at the later stage of growth.

A single gene had a decrease in the fold regulation at the later stage of growth when compared to the earlier stage.

These particular genes may be directly regulated by *agr* genes in a cell density dependent manner. Further studies are needed to investigate the true mechanism by which the *agr* system regulates genes in *C.acetobutylicum*. These could involve the use of expression and reporter based vectors and the identification of the AIP. The isolation of the AIP could also be used to complement the mutant.

5.4.5 Less stringent statistical analysis reveals supporting evidence that the *agr* system regulates sporulation and solventogenesis

This chapter has focused on those genes which have been significantly up-regulated in the microarray. That is that the genes have been analysed using the appropriate statistical algorithm in order to provide robust evidence of the *agr* systems potential regulatory role. This has revealed potential mechanisms by which an *agr* system may function in *C.acetobutylicum*.

The data presented in this chapter have been filtered with T-test using a p-value cut-off of 0.01. A multiple testing correction (MTC) was also applied, the Benjamini and Hochberg False Discovery Rate (BHFDR). This reduces the number of genes which could have passed through the t-test by chance. Furthermore the 10 % least statistically robust of genes were removed by standard deviation

When no BHFDR is applied, a greater number of genes can be seen to have differential expression. Although this variation is not as statistically robust, some of the genes add weight to the evidence towards a proposed model of *agr* regulation in *C.acetobutylicum*.

This can be seen in the next chapter in figure 6.1, which shows some of the genes which were found to be down regulated in the *agrB* mutant. This provides further evidence to support a possible model by which the *agr* system regulates sporulation.

6.0 General Discussion

6.1 Regulation of the *agr* homologues in *C.acetobutylicum*

The initial part of this study looked at the transcriptional linkage of the *agr* homologues in *C.acetobutylicum*. The RT-PCR used found linkage between the *agrB* and *agrD* ORF and the *agrC* and *agrA* ORF's. Some evidence was obtained using RNA isolated from late exponential phase cells that *agrD* and *agrC* ORF's were transcriptionally linked. This was most likely due to an aberrant PCR product. The fact that the reporter assays revealed that *agrC* was up-regulated during the late exponential phase of growth, and that *agrB* was not provides evidence that this is the case.

Both the RT-PCR and reporter assay data suggest that *agrC* and *agrA* form part of a single operon and additionally show that *agrB* and *agrD* are also linked. The reporter assay data suggest that *agrB* has its own promoter distinct from *agrC* as the CAT activity obtained with the *agrB*::*catP* construct was dissimilar to that obtained with the *agrC*::*catP* construct. This contradicts previous microarray data which suggests that expression of *agrB* follows the established paradigm for *agrB* expression in *S.aureus* by being up-regulated during stationary phase (Jones *et al.,* 2008).

It is difficult to elucidate why the CAT activity was so limited with respect to the *agrB* promoter based vector. The vector was checked for sequence errors twice by sequencing and restriction analysis. The aberration could have been caused by a mutation within the CAT gene itself after transformation which caused a reduction in its acetyl transferase activity. With respect to the whole *agr* region *agrB* was naturally found to be up-regulated 72-fold at the early stage of growth and 1274-fold at the later stage of growth in the wild type compared to the *agrB* mutant.

The same pattern occurred with the *agrD* homologue which was up-regulated 129-fold at the early stage of growth and 96762-fold at the later stage of growth. The same was not true for *agrC* and *agrA*. *agrC* was only slightly up-regulated 4.17-fold at the early stage of growth and the *agrA* homologue was not differentially expressed at either stage. When the t-test was not used there was some difference at the later stage of growth but the fold change was not found to be as great as with the *agrB* and *agrD* genes. This does lend some further evidence to suggest the *agrB/D* and *agrC/A* regions are not linked. Further evidence is needed to see if these genes are part of the same regulon.

6.2 Generation of *agr* knockout mutants

Many problems are encountered which had to be solved when undertaking genetic studies in *C.acetobutylicum*. Along with many other members of the genus *C.acetobutylicum* is a notoriously difficult organism to work with and can grow inconsistently. In batch culture *C.acetobutylicum* was initially found to grow most consistently in Clostridial Basal Medium (CBM). The use of MES to buffer the medium in further studies helped with growth consistency. *C.acetobutylicum* is sometimes cited as one of the most amenable of the clostridia to genetic manipulation.

Personal communications from other groups have revealed subtly different methods to transform this organism. These range from electroporating the organism at different time points in the growth curve, varying the medium used to grow the organism, and electroporating in an aerobic environment rather than in an anaerobic cabinet. Efforts were made to try and improve the frequency obtained using existing methods. For example, the use of square waves based on published evidence (Tyurin *et al*., 2000) was employed with no success.

The use of conjugation to transform *C.acetobutylicum* was investigated with some success but this was not deemed suitable as it was no more reproducible than electroporation and is more time consuming. The most

reproducible method was found to be based on a published method (Mermelstein and Papoutsakis, 1993). This method proved adequate for the genetic manipulations needed for production of most of the mutants obtained and was more than sufficient to obtain all the ClosTron mutants. This uses cells grown to an OD_{600} of 0.9. A preliminary experiment was conducted to see if it would be possible to transform cells which had been previously grown to this stage and stored in glycerol at -80ºC for a month with a pMTL007 based plasmid. Transformants were obtained, albeit with slightly lower efficiency. No further investigation was done to see if mutants could be made as the required mutants had already been obtained but the use of glycerol stocks of electrocompetent *C.acetobutylicum* could reduce the amount of time it takes to get gene knockouts in future studies.

The generation of knockout mutants in *C.acetobutylicum* represented the largest component of this study. The development of the antisense RNA vectors was initially the only possible procedure available for obtaining gene disruption in this organism. Four antisense vectors were made and RT-PCR did indicate an antisense effect on the gene products of the first two vectors pJS06 and pJS07.

Before further investigations into the use of antisense RNA vectors and development of suitable phenotypic assays could progress, a method of obtaining full gene knockout became available. The ability to generate targeted gene disruption in this important organism will greatly improve future studies. Using the vector based method developed in the laboratory of Phillipe Soucaille it was possible to obtain a knockout of *agrA* by allelic exchange.

The fusion PCR method used by this laboratory although relatively quick did not seem to produce accurate sequences of the flanking regions of *agrA*. This was probably due to the fact that the length of the template was so long that even a high proof reading Taq polymerase was unable to read it perfectly. The answer was to obtain individual copies of the flanking

regions by PCR which because they were smaller reduced the chance of a misread. It was then possible to ligate these regions in a three way ligation with the MLS cassette and the pCLF.2.1 vector to get the desired vector.

This method was found to be the only method in which it was possible to get the desired vector with no misreads. The splicing out of the MLS cassette was achieved only at the lower temperature of 30ºC which has implications for laboratories working with heated anaerobic cabinets. Some laboratories work with *C.acetobutylicum* at ambient temperatures where this problem does not occur.

Although this method makes it possible to get a complete deletion of the target gene it was found to take a great deal of time and effort to clone the desired vector and obtain a mutant. The development of a Targetron based system by John Heap has made it possible to obtain gene knockout in a wide variety of different clostridial species (Heap *et al.,* 2007). Using this system it was possible to obtain 6 different mutant strains of *C.acetobutylicum* in the same time it took to obtain the single *agrA* whole gene knockout. Three genes in the putative *agr* system were targeted including *agrB*, *agrC* and *agrA*. Unfortunately due to the small size of the *agrD* homologue it was not possible to obtain a target site for this system. Furthermore the inclusion of FRT sites in the RAM of the pMTL007C-E2 ClosTron vectors allowed for its excision with pCLF-1, leaving a “clean” insertional mutation.

Given the fact that so many mutants were obtained in this study it is a pity that they could not be studied in greater detail. This was mainly due to time constraints as a lot of time was needed in attempting to obtain the mutants. The original three *agr* mutants with the RAM were used first in growth curves to look at product formation as a method for excising the MLS cassette in the whole *agrA* mutant was still being developed. When it became available it was then applied to the second generation ClosTron mutants with FRT sites. Once this technique was perfected the three second generation *agr* mutants and the *agrA* deletion mutant where used

in growth curves to look at product formation. The development of two ways to achieve whole chromosomal knockout and targeted gene insertion is a great breakthrough in the study of *C.acetobutylicum*. Other clostridial species have also been successfully targeted using ClosTron technology (Heap *et al*., 2007). The ability to obtain gene knockout will inevitably increase the knowledge of functional genomics in this genus. More recently the ClosTron technology has been described in publications were it has been used to target genes such as *cwpV* in *C.difficile* (Emerson *et al.,* 2009), *spo0A* and a previously uncharacterised histidine kinase in *C.difficile* (Underwood *et al.,* 2009). In the latter study the knockouts were found to have reduced toxin production. The ClosTron has also been used to knockout *acp*, a peptidoglycan hydrolase of *C.perfringens* and to produce a non-toxic strain of *C.botulinum* for use in food research (Camiade *et al*., 2010; Bradshaw *et al.*, 2010). In *C.acetobutylicum* this system has been used to knockout the type II restriction system avoiding the need to methylate vectors before transformation into this strain (Dong *et al.,* 2010).

6.3 Phenotypic effects of *agr* mutation

The ability to produce so many mutants in so short a space of time led to the problem of studying the phenotypic effect of the mutations. In pathogenic clostridia the most obvious phenotype to look at is the production of toxins and there are many readily available assays to do this. Because *C.acetobutylicum* is a commercially important organism the first phenotypic effect of mutation to examine was logically the production of acids and solvents in the mutants. There was no simple assay to look at this as HPLC and gas chromatography are the most commonly used methods to look at the levels of acids and solvents in the medium. This study used gas chromatography methods to analyse the levels of product formation in the *agr* mutants compared to those of the wild type. The levels of most products were found not to be significantly different from wild type levels with the majority of the *agr* mutants. These results were affected by the medium used which did not allow the full production of

products by *C.acetobutylicum*. It is tempting to suggest that the *agr* system has no effect on solvent production and indeed this would be a useful finding as the evidence from this study also suggests that *agr* has a major effect on sporulation. A recent study has using mutants of *agr* have also found an effect on sporulation but no effect on solventogenesis (Steiner *et al*., 2012). A mutation which reduces the sporulation event in this organism yet has no effect on solventogenesis would be beneficial commercially. If the *agr* system has an effect on sporulation by interacting with the phosphorylation of Spo0A but has no effect on solventogenesis It would not take into account the fact that Spo0A is thought to be implicated in the switch from acidogenesis to solventogenesis. It is therefore unfortunate that the effect on the mutation of *agr* on solventogenesis was found to be inconclusive. If the mutation of *agr* was found to lower the amount of solvents then it would fit in perfectly with this model.

6.4 Future studies

The main body of work in this study was directed into obtaining gene disruption in *C.acetobutylicum*. It is a testament to the rapid advances in this area that it was possible to obtain two different types of mutant. The large amount of time needed to obtain these knockouts meant that there was limited time left to develop the analytical methods needed to look at the effects of gene disruption of this organism.

Most published studies of this organism involve growing it in a bioreactor. It has only been possible to grow the organism in a normal batch culture at 37ºC. For this reason it has proved difficult to compare the findings of this study with existing published studies.

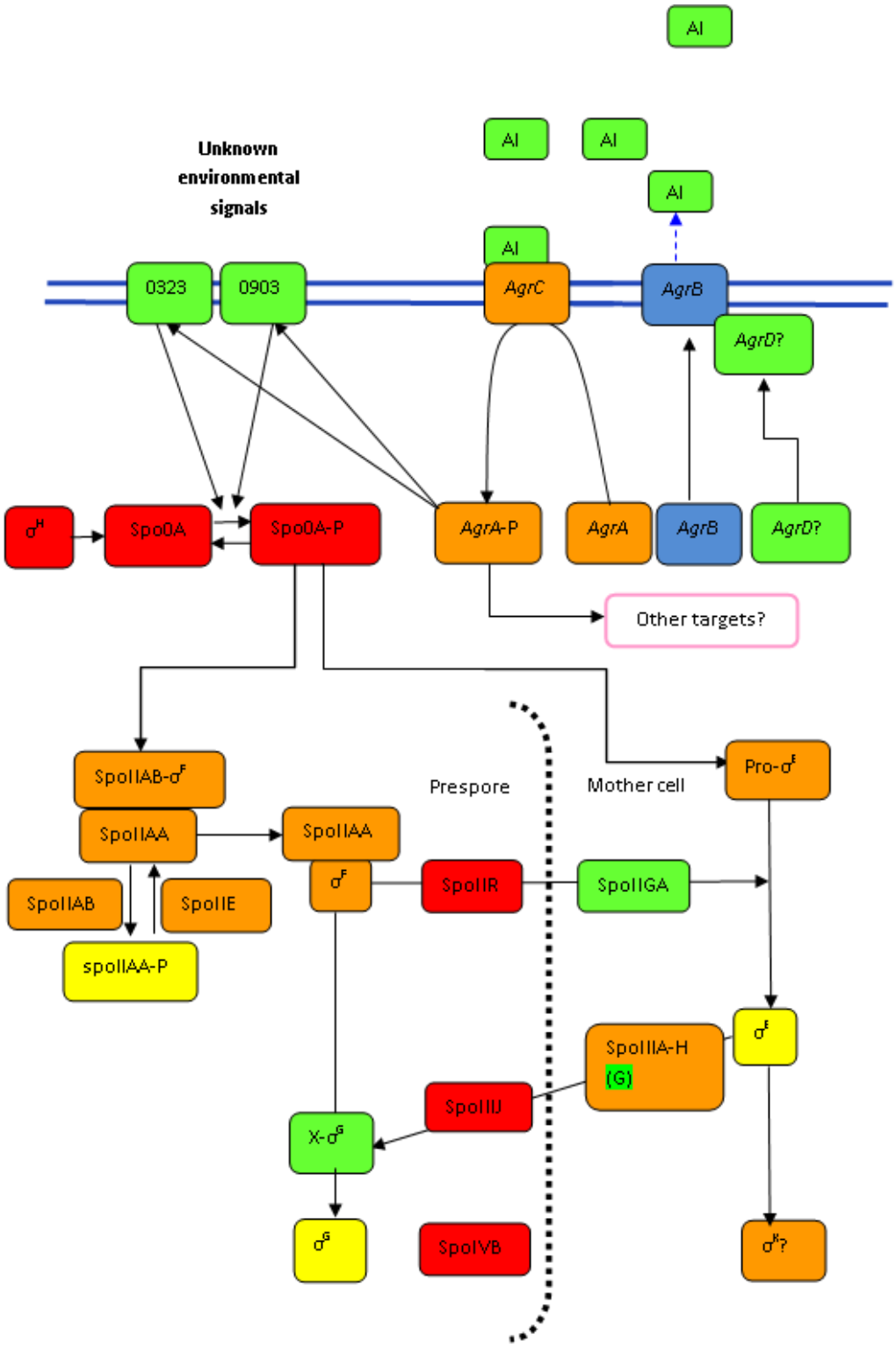


Figure 6.1 Model showing the possible mechanism by which the *agr* system in *C.acetobutylicum* regulates sporulation via quorum sensing. Those genes found to be significantly down-regulated in the *agrB* mutant are represented by green boxes (this also applies with SpoIIIAG). Those genes found to be down-regulated in the *agrB* mutant when less stringent statistical methods were applied are shown in orange boxes. Those found to be unaffected by *agrB* are shown as red boxes. Yellow boxes represent post transcriptional modifications. This model is based on the model used for *B.subtilis* and may have significant differences in *C.acetobutylicum*. The existence of a true σ^K is still undetermined.

It is probably the case that there is differential gene expression in the organism grown in normal batch culture conditions compared to when it is grown in a buffered bioreactor.

This study has produced several mutants of the components in the putative *agr* system in *C.acetobutylicum*. It has revealed that knocking out the *agr* homologues in this species affects sporulation and knocking out the *agrB* gene has an effect on many genomic and pSOL-1 genes. If less robust statistical analysis of the microarray results is used then it is possible to show the *agr* system being applied to the hypothetical model shown in figure 6.1.

Further evidence is needed to assess whether the putative *agr* system in *C.acetobutylicum* is indeed a quorum sensing-based regulatory system involving all 4 *agr* homologues. A recent study involving knockouts of the two *agr* genes (*agrB* and *agrD*) in *C.perfringens* found that the mutant did not express theta toxin and the transcription of the alpha and kappa toxins was significantly reduced (Ohtani et al., 2009). This enabled the simple experiment of adding the wild type supernatant to the culture to see if the toxin production was restored in effect complementing the mutation by the addition of a potential AIP. This restored toxin production and therefore provided evidence of a potential quorum sensing *agr* system. It is more difficult to repeat this experiment due to the fact there is an absence of a toxin equivalent but this study could provide potential targets to use as an assay to test the hypothesis that there is a true *agr* quorum sensing system in *C.acetobutylicum*. Furthermore, if future studies confirm that the *agr* system does have an effect on product formation then product formation could be used for a similar experiment.

If the *agr* paradigm applies to *C.acetobutylicum* it would be expected that there would be a functional AIP involved. An initial search for the *agrD* product was conducted using mass spectrometry with no success. Further work using the *agrB* and *agrD* mutants may lead to future achievement.

There were many problems in obtaining true complementation of the mutants. Further work needs to be done to ensure that it is possible to obtain true complementation of these mutants. This is particularly relevant due to the potential rapid increase in the number of mutants that can be obtained. The use of other plasmids should be investigated as vectors for complementary genes.

This study raises the possibility that the *agr* system in *C.acetobutylicum* affects sporulation without necessarily affecting solvent formation. This potentially useful phenotype should be examined further under bioreactor conditions.

Bibliography

Abanes-De Mello A, Sun YL, Aung S, Pogliano K (2002). A cytoskeleton-like role for the bacterial cell wall during engulfment of the *Bacillus subtilis* forespore. *Genes Dev*. (24):3253-3264.

:**Abe A, Koide H, Kohno T, Watabe K (1995).** A *Bacillus subtilis* spore coat polypeptide gene, *cotS*. *Microbiology*. 141 (Pt 6):1433-42.

Acetone – PubChem Public Chemical Database
http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=180

:**Alsaker KV, Spitzer TR, Papoutsakis ET (2004).** Transcriptional analysis of *spo0A* overexpression in *Clostridium acetobutylicum* and its effect on the cell's response to butanol stress. *J Bacteriol*. 186(7):1959-1971.

Alsaker KV, Papoutsakis ET (2005). Transcriptional program of early sporulation and stationary-phase events in *Clostridium acetobutylicum*. *J Bacteriol*. 187(20):7103-7118.

Andersch, W, Bahl H, and Gottschalk G (1983). Level of enzymes involved in acetate, butyrate, acetone and butanol formation by *Clostridium acetobutylicum*. *Eur. J. Appl. Microbiol*. Biotechnol. 18:327-332.

Ansaldi M, Marolt D, Stebe T, Mandic-Mulec I, Dubnau D(2002)..Specific activation of the *Bacillus* quorum-sensing systems by isoprenylated pheromone variants. *Mol Microbiol*.;44(6):1561-1573.

Autret N, Raynaud C, Dubail I, Berche P, Charbit A (2003). Identification of the *agr* locus of *Listeria monocytogenes*: role in bacterial virulence. *Infect Immun*. 71(8):4463-4471.

Awad MM, Bryant AE, Stevens DL, Rood JI (1995). Virulence studies on chromosomal alpha-toxin and theta-toxin mutants constructed by allelic exchange provide genetic evidence for the essential role of alpha-toxin in *Clostridium perfringens*-mediated gas gangrene. *Mol Microbiol*. 15(2):191-202.

Awad MM, Ellemor DM, Bryant AE, Matsushita O, Boyd RL, Stevens DL, Emmins JJ, Rood JI (2000). Construction and virulence testing of a collagenase mutant of *Clostridium perfringens*. *Microb Pathog*. 28(2):107-117.

Awad MM, Ellemor DM, Boyd RL, Emmins JJ, Rood JI (2001).. Synergistic effects of alpha-toxin and perfringolysin O in *Clostridium perfringens*-mediated gas gangrene. *Infect Immun*. 69(12):7904-7910.

Awad MM, Rood JI (2002). Perfringolysin O expression in *Clostridium perfringens* is independent of the upstream *pfoR* gene. *J Bacteriol*. 184(7:2034-2038.

Bacon Schneider K, Palmer TM, Grossman AD. (2002). Characterization of *comQ* and *comX*, two genes required for production of ComX pheromone in *Bacillus subtilis*. *J Bacteriol*. 184(2):410-419.

Bagyan I, Hobot J, Cutting S (1996). A compartmentalized regulator of developmental gene expression in *Bacillus subtilis*. *J Bacteriol*. 178(15):4500-4507.

Bai U, Mandic-Mulec I, Smith I (1993). SinI modulates the activity of SinR, a developmental switch protein of *Bacillus subtilis*, by protein–protein interaction. *Genes Dev* 7: 139–148.

Bannam TL, Crellin PK, Rood JI. (1995). Molecular genetics of the chloramphenicol-resistance transposon Tn*4451* from *Clostridium perfringens*: the TnpX site-specific recombinase excises a circular transposon molecule. *Mol Microbiol*. May;16(3):535-551.

Banu S, Ohtani K, Yaguchi H, Swe T, Cole ST, Hayashi H, Shimizu T (2000). Identification of novel VirR/VirS-regulated genes in *Clostridium perfringens*. *Mol Microbiol*. 35(4):854-864.

Bassler BL, Wright M, Showalter RE, Silverman MR (1993). Intercellular signalling in *Vibrio harveyi*: sequence and function of genes regulating expression of luminescence. *Mol Microbiol*. 1993 Aug;9(4):773-786.

Bates S, Roscoe RA, Althorpe NJ, Brammar WJ, Wilkins BM (1999). Expression of leading region genes on IncI1 plasmid ColIb-P9: genetic evidence for single-stranded DNA transcription. *Microbiology*. 145 (10):2655-2662.

Bath J, Wu LJ, Errington J, Wang JC (2000). Role of *Bacillus subtilis* SpoIIIE in DNA transport across the mother cell-prespore division septum. *Science*. 290(5493):995-997.

Bennett, GN, Rudolph, FB (1995). The central metabolic pathway from acetyl-CoA to butyryl-CoA in *Clostridium acetobutylicum*. *FEMS Microbiol Rev* 17, 241–249.

Ben-Yehuda S, Losick R (2002). Asymmetric cell division in *B. subtilis* involves a spiral-like intermediate of the cytokinetic protein FtsZ. *Cell*.109(2):257-266.

Ben-Yehuda S, Rudner DZ, Losick R (2003a). Assembly of the SpoIIIE DNA translocase depends on chromosome trapping in *Bacillus subtilis*. *Curr Biol*. 16;13(24):2196-2200.

Ben-Yehuda S, Rudner DZ, Losick R (2003b). RacA, a bacterial protein that anchors chromosomes to the cell poles. *Science*. 299(5606):532-536.

Bertram, J. Dürre, P (1989). Conjugal transfer and expression of streptococcal transposons in *Clostridium acetobutylicum*. *Arch Microbiol*. 151, 551-557.

Bischofs IB, Hug JA, Liu AW, Wolf DM, Arkin AP (2009). Complexity in bacterial cell-cell communication: quorum signal integration and subpopulation signaling in the *Bacillus subtilis* phosphorelay.*Proc Natl Acad Sci U S A*. 21;106(16):6459-6464.

Bjorklind A, Arvidson S (1980). Mutants of *Staphylococcus aureus* affected in the regulation of exoprotein synthesis. *FEMS Microbiology Letters* 7,203-206.

Bogush M, Xenopoulos P, Piggot PJ (2007). Separation of chromosome termini during sporulation of *Bacillus subtilis* depends on SpoIIIE. *J Bacteriol*. 189(9):3564-3572.

Borden JR, Papoutsakis ET (2007). Dynamics of genomic-library enrichment and identification of solvent tolerance genes for *Clostridium acetobutylicum. Appl Environ Microbiol*. 73(9):3061-3068.

Boukhvalova M, VanBruggen R, Stewart RC (2002). CheA kinase and chemoreceptor interaction surfaces on CheW. *J Biol Chem*. 277(26):23596-23603.

Bourgogne A, Hilsenbeck SG, Dunny GM, Murray BE. (2006). Comparison of OG1RF and an isogenic *fsrB* deletion mutant by transcriptional analysis: the Fsr system of *Enterococcus faecalis* is more than the activator of gelatinase and serine protease. *J Bacteriol*. 188(8):2875-2884.

Bowles LK, Ellefson WL (1985). Effects of butanol on *Clostridium acetobutylicum*. *Appl Environ Microbiol*. 50(5):1165-1170.

Boynton Z L, Bennett GN, Rudolph FB (1996). Cloning, sequencing, and expression of clustered genes encoding β-hydroxybutyryl-coenzyme A (CoA). dehydrogenase, crotonase, and butyryl-CoA dehydrogenase from *Clostridium acetobutylicum* ATCC 824. *J Bacteriol* 178, 3015–3024.

Bradshaw M, Marshall KM, Heap JT, Tepp WH, Minton NP, Johnson EA (2010). Construction of a nontoxigenic *Clostridium botulinum* strain for food challenge studies. *Appl Environ Microbiol*. 76(2):387-393.

Beenken KE, Mrak LN, Griffin LM, Zielinska AK, Shaw LN, Rice KC, Horswill AR, Bayles KW, Smeltzer MS (2010). Epistatic relationships between *sarA* and *agr* in *Staphylococcus aureus* biofilm formation. *PLoS One*. May 24;5(5):e10790.

Boles BR, Horswill AR (2008). Agr-mediated dispersal of *Staphylococcus aureus* biofilms. *PLoS Pathog*. Apr 25;4(4):e1000052.

Broder DH, Pogliano K (2006). Forespore engulfment mediated by a ratchet-like mechanism. *Cell*. 8;126(5):917-28.

Bryant DL, Blaschek HP (1988).. Buffering as a means for increasing growth and butanol production by *Clostridium acetobutylicum*. *J Ind Microbiol Biotechnol*.3:49–55.

Buchholz F, Ringrose L, Angrand P, Rossi F, Stewart A (1996). Different thermostabilities of FLP and Cre recombinases: implications for applied site-specific recombination. *Nucleic Acids Research* 24 (21). 4256–4262.

Bullifent HL, Moir A, Titball RW (1995). The construction of a reporter system and use for the investigation of *Clostridium perfringens* gene expression. *FEMS Microbiol Lett*. (1):99-105.

Camiade E, Peltier J, Bourgeois I, Couture-Tosi E, Courtin P, Antunes A, Chapot-Chartier MP, Dupuy B, Pons JL (2010). Characterization of Acp, a peptidoglycan hydrolase of *Clostridium perfringens* with N-acetylglucosaminidase activity, implicated in cell separation and stress-induced autolysis. *J Bacteriol*. 192(9):2373-2384.

Cámara M, Williams P, Hardman A (2002). Controlling infection by tuning in and turning down the volume of bacterial small-talk. *Lancet Infect Dis*. Nov;2(11):667-676.**Camp AH, Losick R (2008)**. A novel pathway of intercellular signalling in *Bacillus subtilis* involves a protein with similarity to a component of type III secretion channels. *Mol Microbiol*. 69(2):402-417.

Camp AH, Losick R. (2009). A feeding tube model for activation of a cell-specific transcription factor during sporulation in *Bacillus subtilis*. *Genes Dev*. Apr 15;23(8):1014-1024.

Carter GP, Purdy D, Williams P, Minton NP (2005). Quorum sensing in *Clostridium difficile*: analysis of a *luxS*-type signalling system. *J Med Microbiol*. 54(2):119-127.

Carter Glen (2004). The Development of Molecular Tools and the Study of LuxS in *Clostridium difficile*. Ph.D. Thesis. University of Nottingham: UK

Cary JW, Petersen DJ, Papoutsakis ET, Bennett GN (1988). Cloning and expression of *Clostridium acetobutylicum* phosphotransbutyrylase and butyrate kinase genes in *Escherichia coli*. *J Bacteriol*. 170(10):4613–4618.

Cervin MA, Lewis RJ, Brannigan, JA, and Spiegelman GB (1998). The *Bacillus subtilis* regulator SinR inhibits *spoIIG* promoter transcription in vitro without displacing RNA polymerase. *Nucleic Acids Res* 26: 3806–3812.

Cha JH, Stewart GC (1997). The *divIVA* mini cell locus of *Bacillus subtilis*. *J Bacteriol*. 179(5):1671-1683.

Chan WC, Coyle BJ, Williams P (2004). Virulence regulation and quorum sensing in *staphylococcal* infections: competitive AgrC antagonists as quorum sensing inhibitors. *J Med Chem*. 47(19):4633-4641.

Chary VK, Meloni M, Hilbert DW, Piggot PJ (2005). Control of the expression and compartmentalization of (sigma).G activity during sporulation of *Bacillus subtilis* by regulators of (sigma).F and (sigma).E. *J Bacteriol*. 187(19):6832-6840.

Chary VK, Xenopoulos P, Piggot PJ.(2006). Blocking chromosome translocation during sporulation of *Bacillus subtilis* can result in prespore-specific activation of sigmaG that is independent of sigmaE and of engulfment. *J Bacteriol*. 188(20):7267-7273.

Chary VK, Xenopoulos P, Piggot PJ (2007). Expression of the sigmaF-directed *csfB* locus prevents premature appearance of sigmaG activity during sporulation of *Bacillus subtilis*. *J Bacteriol*. 189(23):8754-8757.

Chen Y, McClane BA, Fisher DJ, Rood JI, Gupta P (2005). Construction of an alpha toxin gene knockout mutant of *Clostridium perfringens* type A by use of a mobile group II intron. *Appl Environ Microbiol*. 71(11):7542-7547.

Chugani SA, Whiteley M, Lee KM, D'Argenio D, Manoil C, Greenberg EP (2001). QscR, a modulator of quorum-sensing signal synthesis and virulence in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A*. Feb 27;98(5):2752-7.

Collett HJ, Woods DR Reid, SJ (1997). Characterisation of a transposon-induced pleiotropic mutant of *Clostridium acetobutylicum* P262. *Anaerobe* 3(6). 411-421.

Comella N, Grossman AD (2005). Conservation of genes and processes controlled by the quorum response in bacteria: characterization of genes controlled by the quorum-sensing transcription factor ComA in *Bacillus subtilis*. *Mol Microbiol*. 57(4):1159-1174.

Cooksley CM, Davis IJ, Winzer K,. Chan WC, Peck MW and Minton NP (2010). Regulation of neurotoxin production and sporulation by a Putative *agrBD* signaling system in proteolytic *Clostridium botulinum*. *Appl Environ Microbiol*. 2010 76(13):4448-4460.

Coote, JG. (1972). Sporulation in *Bacillus subtilis,* Characterization of oligosporogenous mutants and comparison of their phenotypes with those of asporogenous mutants. *J. Gen. Microbiol*. 71:1-15.

Core LJ, Perego M. (2003). TPR-mediated interaction of RapC with ComA inhibits response regulator-DNA binding for competence development in *Bacillus subtilis*. Mol. Microbiol. 49:1509–1522.

Cornillot E, Nair RV, Papoutsakis ET, Soucaille P (1997). The genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 reside on a large plasmid whose loss leads to degeneration of the strain. *J Bacteriol*. 179:5442-5447

Cousineau B, Lawrence S, Smith D, Belfort M (2000). Retrotransposition of a bacterial group II intron. *Nature*. 404(6781):1018-1021.

Cousineau B, Smith D, Lawrence-Cavanagh S, Mueller JE, Yang J, Mills D, Manias D, Dunny G, Lambowitz AM, Belfort M (1998). Retrohoming of a bacterial group II intron: mobility via complete reverse splicing, independent of homologous DNA recombination. *Cell*. 94(4):451-462.

Daniel RA, Drake S, Buchanan CE, Scholle R, Errington J (1994). The *Bacillus subtilis spoVD* gene encodes a mother-cell-specific penicillin-binding protein required for spore morphogenesis. *J Mol Biol*. 235(1):209-220.

de Boer PA, Crossley RE, Hand AR, Rothfield LI (1991). The MinD protein is a membrane ATPase required for the correct placement of the *Escherichia coli* division site. *EMBO J*. 10(13):4371-4380.

Deziel E, Gopalan S, Tampakaki AP, Lepine F, Padfield KE, Saucier M, Xiao G, Rahme LG (2005). The contribution of MvfR to *Pseudomonas aeruginosa* pathogenesis and quorum sensing circuitry regulation: multiple quorum sensing-regulated genes are modulated without affecting *lasRI*, *rhlRI* or the production of N-acyl-L-homoserine lactones. *Mol.Microbiol*. 55:998–1014.

Delver EP, Kotova VU, Zavilgelsky GB, Belogurov AA. (1991). Nucleotide sequence of the gene (*ard*). encoding the antirestriction protein of plasmid colIb-P9. *J Bacteriol*. 173(18):5887-92.

Desai RP, Harris LM, Welker NE, Papoutsakis ET (1999). Metabolic flux analysis elucidates the importance of the acid-formation pathways in regulating solvent production by *Clostridium acetobutylicum*. Metab Eng. 1(3):206-213.

Diggle SP, Winzer K, Chhabra SR, Worrall KE, Camara M, Williams P (2003). The *Pseudomonas aeruginosa* quinolone signal molecule overcomes the cell density-dependency of the quorum sensing hierarchy, regulates *rhl*-dependent genes at the onset of stationary phase and can be produced in the absence of LasR. *Mol. Microbiol*. 50:29–43.

Dineen SS, Villapakkam AC, Nordman JT, Sonenshein AL (2007). Repression of *Clostridium difficile* toxin gene expression by CodY. *Mol Microbiol*. 66(1):206-219.

Doan T, Marquis KA, Rudner DZ (2005). Subcellular localization of a sporulation membrane protein is achieved through a network of interactions along and across the septum. *Mol Microbiol*. 55(6):1767-1781.

Doan T, Morlot C, Meisner J, Serrano M, Henriques AO, Moran CP Jr, Rudner DZ (2009). Novel secretion apparatus maintains spore integrity and developmental gene expression in *Bacillus subtilis*. *PLoS Genet*. 5(7):e1000566. Epub 2009 Jul 17.

Dong H, Zhang Y, Dai Z, Li Y (2010). Engineering *clostridium* strain to accept unmethylated DNA. *PLoS One*. 5(2):e9038

Dubey AK, Baker CS, Suzuki K, Jones AD, Pandit P, Romeo T, Babitzke P (2003). CsrA regulates translation of the *Escherichia coli* carbon starvation gene, *cstA*, by blocking ribosome access to the *cstA* transcript. *J Bacteriol*. 185(15):4450-4460.

Dubnau D. Lovett CM, Jr. (2002). in *Bacillus subtilis* and Its Closest Relatives: From Genes to Cells, eds. Sonenshein, A. L., Hoch, J. A. & Losick, R. (Am. Soc. Microbiol., Washington, DC)., pp. 453–471.

Duncan L, Alper S, Arigoni F, Losick R, Stragier P (1995). Activation of cell-specific transcription by a serine phosphatase at the site of asymmetric division. *Science*. 270(5236):641-644.

Dunny GM, Leonard BA (1997). Cell-cell communication in gram-positive bacteria. *Annu Rev Microbiol*.51:527-564.

Dürre P (2007). Biobutanol: an attractive biofuel. *Biotechnol J*. 2(12):1525-1534.

Edwards DH, Errington J (1997). The *Bacillus subtilis* DivIVA protein targets to the division septum and controls the site specificity of cell division. *Mol Microbiol*. 24(5):905-915.

Eichenberger P, Fawcett P, Losick R (2001). A three-protein inhibitor of polar septation during sporulation in *Bacillus subtilis*. *Mol Microbiol*. 42(5):1147-1162.

Eichenberger P, Fujita M, Jensen ST, Conlon EM, Rudner DZ, Wang ST, Ferguson C, Haga K, Sato T, Liu JS, Losick R (2004). The program of gene transcription for a single differentiating cell type during sporulation in *Bacillus subtilis*. *PLoS Biol*. 2(10):e328.

Eldakak A, Hulett FM (2007). Cys303 in the histidine kinase PhoR is crucial for the phosphotransfer reaction in the PhoPR two-component system in *Bacillus subtilis*. *J Bacteriol*. 189(2):410-421.

Emerson JE, Reynolds CB, Fagan RP, Shaw HA, Goulding D, Fairweather NF. (2009). A novel genetic switch controls phase variable expression of CwpV, a *Clostridium difficile* cell wall protein. *Mol Microbiol*. Nov;74(3):541-556. **Errington J, Mandelstam J. (1984).** Genetic and phenotypic characterization of a cluster of mutations in the *spoVA* locus of *Bacillus subtilis*. *J Gen Microbiol*. 130(8):2115-2121.

Errington J, Appleby L, Daniel RA, Goodfellow H, Partridge SR, Yudkin MD (1992). Structure and function of the *spoIIIJ* gene of *Bacillus subtilis*: a vegetatively expressed gene that is essential for sigma G activity at an intermediate stage of sporulation. *J Gen Microbiol*. 138(12):2609-2618.

Ezeji TC, Qureshi N, Blaschek HP (2007). Bioproduction of butanol from biomass: from genes to bioreactors.*Curr Opin Biotechnol*. 18(3):220-227.

Fan N, Cutting S, Losick R (1992). Characterization of the *Bacillus subtilis* sporulation gene *spoVK*. *J Bacteriol*. 174(3):1053-1054.

Ferrari FA, Trach K, LeCoq D, Spence J, Ferrari E, Hoch JA. (1985). Characterization of the *spo0A* locus and its deduced product. *Proc Natl Acad Sci U S A*. 82(9):2647-51.

Feustel L, Nakotte S, Dürre P (2004). Characterization and development of two reporter gene systems for *Clostridium acetobutylicum*. *Appl Environ Microbiol*. 70(2):798-803.

Fierro-Monti IP, Reid SJ, Woods DR (1992). Differential expression of a *Clostridium acetobutylicum* antisense RNA: implications for regulation of glutamine synthetase. *J Bacteriol*. 174(23):7642-7647.

Fisher, SH, Strauch, MA, Atkinson, MR, and Wray, LV, Jr (1994). Modulation of *Bacillus subtilis* catabolite repression by transition state regulatory protein AbrB. *J. Bacteriol*. 176: 1903–1912.

Fontaine L, Meynial-Salles I, Girbal L, Yang X, Croux C, Soucaille P (2002). Molecular characterization and transcriptional analysis of *adhE2*, the gene encoding the NADH-dependent aldehyde/alcohol dehydrogenase responsible for butanol production in alcohologenic cultures of *Clostridium acetobutylicum* ATCC 824. *J Bacteriol*. 184(3):821-830.

Frazier CL, San Filippo J, Lambowitz AM, Mills DA (2003). Genetic manipulation of *Lactococcus lactis* by using targeted group II introns: generation of stable insertions without selection. *Appl Environ Microbiol*. 69(2):1121-1128.

Fujii T, Ingham C, Nakayama J, Beerthuyzen M, Kunuki R, Molenaar D, Sturme M, Vaughan E, Kleerebezem M, de Vos W. (2008). Two homologous Agr-like quorum-sensing systems cooperatively control adherence, cell morphology, and cell viability properties in *Lactobacillus plantarum* WCFS1**.** *J Bacteriol*. 190(23): 7655-7665

Fuqua WC, Winans SC, Greenberg EP (1994). Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *J Bacteriol*. 1994 176(2):269-275.

Gallagher L A, McKnight SL, Kuznetsova MS, Pesci EC, Manoil C (2002) Functions required for extracellular quinolone signaling by *Pseudomonas aeruginosa*. *J. Bacteriol*. 184:6472–6480.

Garmyn D, Gal L, Lemaitre JP, Hartmann A, Piveteau P. (2009). Communication and autoinduction in the species *Listeria monocytogenes*: A central role for the *agr* system. *Commun Integr Biol*. Jul;2(4):371-314.

Gerischer U, Dürre P (1990). Cloning, sequencing, and molecular analysis of the acetoacetate decarboxylase gene region from *Clostridium acetobutylicum*. *J Bacteriol*. 172(12):6907-6918.

Gholamhoseinian A, Shen Z, Wu JJ, Piggot P (1992). Regulation of transcription of the cell division gene *ftsA* during sporulation of *Bacillus subtilis*. *J Bacteriol*. 174(14):4647-4656.

Girbal L, Mortier-Barrière I, Raynaud F, Rouanet C, Croux C, Soucaille P (2003). Development of a sensitive gene expression reporter system and an inducible promoter-repressor system for *Clostridium acetobutylicum*. *Appl Environ Microbiol*. 69(8):4985-4988.

Gonzy-Tréboul G, Karmazyn-Campelli C, Stragier P (1992). Developmental regulation of transcription of the *Bacillus subtilis ftsAZ* operon. *J Mol Biol*. 224(4):967-979.

Graumann PL, Losick R (2001). Coupling of asymmetric division to polar placement of replication origin regions in *Bacillus subtilis*. *J Bacteriol*. 183(13):4052-4060.

Grebe TW, Stock JB. (1999). The histidine protein kinase superfamily. *Adv Microb Physiol*.;41:139-227.

Green EM, Bennett GN (1996). Inactivation of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824. *Appl Biochem Biotechnol*. 57:213-21.

Green EM, Boynton ZL, Harris LM, Rudolph FB, Papoutsakis ET, Bennett GN (1996). Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. *Microbiology*. 142 (8):2079-2086.

Grossman, AD (1995). Genetic networks controlling the initiation of sporulation and the development of genetic competence in *Bacillus subtilis*. *Annu Rev Genet*. 29:477-508.

Guillot C, Moran CP Jr (2007). Essential internal promoter in the *spoIIIA* locus of *Bacillus subtilis*. *J Bacteriol* 189: 7181–7189.

Gutierrez NA, Maddox IS (1987). role of chemotaxis in solvent production by *Clostridium acetobutylicum Appl Environ Microbiol*. 53(8):1924-1927.

Hächler H, Kayser FH, Berger-Bächi B (1987). Homology of a transferable tetracycline resistance determinant of *Clostridium difficile* with *Streptococcus* (*Enterococcus*). *faecalis* transposon Tn*916*. *Antimicrob Agents Chemother*. 31(7):1033-1038.

Halberg R, Kroos L (1994). Sporulation regulatory protein SpoIIID from *Bacillus subtilis* activates and represses transcription by both mother-cell-specific forms of RNA polymerase. *J Mol Biol*. 243(3):425-436.

Hahn, J, Roggiani, M, and Dubnau, D. (1995). The major role of Spo0A in genetic competence is to downregulate *abrB*, an essential competence gene. *J Bacteriol* 177: 3601–3605.

Hancock LE, Perego M. (2004). The *Enterococcus faecalis fsr* two-component system controls biofilm development through production of gelatinase. *J Bacteriol*. 186(17):5629-5639.

Harris LM, Blank L, Desai RP, Welker NE, Papoutsakis ET (2001). Fermentation characterization and flux analysis of recombinant strains of *Clostridium acetobutylicum* with an inactivated *solR* gene. *J Ind Microbiol Biotechnol*. 27(5):322-328.

Harris LM, Welker NE, Papoutsakis ET (2002). Northern, morphological, and fermentation analysis of *spo0A* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. *J Bacteriol*. 184(13):3586-3597.

Hartmanis MG, Gatenbeck S (1984). Intermediary metabolism in *Clostridium acetobutylicum*: levels of enzymes involved in the formation of acetate and butyrate. *Appl Environ Microbiol*. 47(6):1277-1283.

Hatheway CL (1990). Toxigenic clostridia. *Clin Microbiol Rev*. 3(1):66-98.

Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP (2007). The ClosTron: a universal gene knock-out system for the genus *Clostridium*. *J Microbiol Methods. 70(3):*452-564.

Heap JT, Kuehne SA, Ehsaan M, Cartman ST, Cooksley CM, Scott JC, Minton NP.(2010). The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. *J Microbiol Methods.* 80(1):49-55.

Henriques AO, Glaser P, Piggot PJ, Moran CP Jr (1998). Control of cell shape and elongation by the *rodA* gene in *Bacillus subtilis*. *Mol Microbiol*. 28(2):235-247.

Higashi Y, Chazono M, Inoue K, Yanagase Y, Amano T (1973). Complementation of theta-toxinogenicity between mutants of two groups of *Clostridium perfringens*. *Biken J*. 16(1):1-9.

Hilbert DW, Piggot PJ (2004). Compartmentalization of gene expression during *Bacillus subtilis* spore formation. *Microbiol Mol Biol Rev*. 68(2):234-262.

Hilbert, D.W., Chary, V.K., and Piggot, P.J. (2004). Contrasting effects of sigmaE on compartmentalization of sigmaF activity during sporulation of *Bacillus subtilis*. *J. Bacteriol*. 186, 1983–1990.

Hoch JA (1993a). Regulation of the phosphorelay and the initiation of sporulation in *Bacillus subtilis*. *Annu Rev Microbiol*. 47:441-465. **Hoch JA (1993b).** The phosphorelay signal transduction pathway in the initiation of *Bacillus subtilis* sporulation. *J Cell Biochem*. 51(1):55-61. **Hu Z,**

Mukherjee A, Pichoff S, Lutkenhaus J (1999). The MinC component of the division site selection system in *Escherichia coli* interacts with FtsZ to prevent polymerization. *Proc Natl Acad Sci U S A*. 96(26):14819-14824.

Huang IH, Waters M, Grau RR, Sarker MR (2004). Disruption of the gene (*spo0A*). encoding sporulation transcription factor blocks endospore formation and enterotoxin production in enterotoxigenic *Clostridium perfringens* type A. *FEMS Microbiol Lett*. 233(2):233-240.

Hussain HA, Roberts AP, Mullany P (2005). Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Deltaerm). and demonstration that the conjugative transposon Tn916DeltaE enters the genome of this strain at multiple sites. *J Med Microbiol*. 54(2):137-141.

Ichikawa H, Halberg R, Kroos L (1999). Negative regulation by the *Bacillus subtilis* GerE protein. *J Biol Chem*. 274(12):8322-8327.

Ichikawa H, Kroos L (2000). Combined action of two transcription factors regulates genes encoding spore coat proteins of *Bacillus subtilis*. *J Biol Chem*. 275(18):13849-13855.

Illing N, Errington J (1991). Genetic regulation of morphogenesis in *Bacillus subtilis*: roles of sigma E and sigma F in prespore engulfment. *J Bacteriol*. 173(10):3159-3169.

Imagawa T, Tatsuki T, Higashi Y, Amano T (1981). Complementation characteristics of newly isolated mutants from two groups of strains of *Clostridium perfringens*. *Biken J*. 24(1-2):13-21.

Imagawa T, Higashi Y (1992). An activity which restores theta toxin activity in some theta toxin-deficient mutants of *Clostridium perfringens*. *Microbiol Immunol*. 36(5):523-527.

Janzon L, Arvidson S (1990). The role of the delta-lysin gene (*hld*). in the regulation of virulence genes by the accessory gene regulator (*agr*). in *Staphylococcus aureus*. *EMBO J*. 9(5):1391-1399.

Ji G, Beavis RC, Novick RP (1995). Cell density control of *staphylococcal* virulence mediated by an octapeptide pheromone. *Proc Natl Acad Sci U S A*. 92(26):12055-12059.

Ji G, Beavis R, Novick RP (1997). Bacterial interference caused by autoinducing peptide variants. *Science*. 276(5321):2027-2030.

Jiang X, Rubio A, Chiba S, Pogliano K (2005). Engulfment-regulated proteolysis of SpoIIQ: evidence that dual checkpoints control sigma activity. *Mol Microbiol*. 58(1):102-115.

Jin S, Levin PA, Matsuno K, Grossman AD, Sonenshein AL. (1997). Deletion of the *Bacillus subtilis* isocitrate dehydrogenase gene causes a block at stage I of sporulation. *J Bacteriol*. Aug;179(15):4725-4732.

Jonas RM, Weaver EA, Kenney TJ, Moran CP Jr, Haldenwang WG (1988). The *Bacillus subtilis spoIIG* operon encodes both sigma E and a gene necessary for sigma E activation. *J Bacteriol*. 170(2):507-511.

Jones DT, Woods DR (1986). Acetone-butanol fermentation revisited. Microbiol Rev. 50 (4). :484-524.

Jones SW, Paredes CJ, Tracy B, Cheng N, Sillers R, Senger RS, Papoutsakis ET (2008). The transcriptional program underlying the physiology of *clostridial* sporulation. *Genome Biol*. 9(7):R114.

Karberg M, Guo H, Zhong J, Coon R, Perutka J, Lambowitz AM (2001). Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. *Nat Biotechnol*. 19(12):1162-1167.

Karmazyn-Campelli C, Rhayat L, Carballido-López R, Duperrier S, Frandsen N, Stragier P (2008). How the early sporulation sigma factor sigmaF delays the switch to late development in *Bacillus subtilis*. *Mol Microbiol*. 67(5):1169-1180. **Karow ML, Glaser P, Piggot PJ (1995).** Identification of a gene, *spoIIR*, that links the activation of sigma E to the transcriptional activity of sigma F during sporulation in *Bacillus subtilis*. *Proc Natl Acad Sci U S A*. 92(6):2012-2016.

Kavanaugh JS, Thoendel M, Horswill AR (2007). A role for type I signal peptidase in *Staphylococcus aureus* quorum sensing. *Mol Microbiol*. 65(3):780-798. **Kellner EM, Decatur A, Moran CP Jr (1996).** Two-stage regulation of an anti-sigma factor determines developmental fate during bacterial endospore formation. *Mol Microbiol*.;21(5):913-924.

Kenney TJ, Moran CP Jr (1987). Organization and regulation of an operon that encodes a sporulation-essential sigma factor in *Bacillus subtilis*. *J Bacteriol*. 169(7):3329-39.

Khvorova A, Zhang L, Higgins ML, Piggot PJ (1998). The *spoIIE* locus is involved in the Spo0A-dependent switch in the location of FtsZ rings in *Bacillus subtilis*. *J Bacteriol*. 180(5):1256-1260.

Kiedrowski MR, Kavanaugh JS, Malone CL, Mootz JM, Voyich JM, Smeltzer MS, Bayles KW, Horswill AR (2011). Nuclease modulates biofilm formation in community-associated methicillin-resistant *Staphylococcus aureus*. *PLoS One*.;6(11):e26714

Kirsch ML, Carpenter PB, Ordal GW (1994). A putative ATP-binding protein from the *che*/*fla* locus of *Bacillus subtilis*. *DNA Seq*. 4(4):271-275.

Kleerebezem M, Quadri LE, Kuipers OP, de Vos WM (1997). Quorum sensing by peptide pheromones and two-component signal-transduction systems in Gram-positive bacteria. *Mol Microbiol*. 24(5):895-904.

Kleerebezem M, Boekhorst J, van Kranenburg R, Molenaar D, Kuipers OP, Leer R, Tarchini R, Peters SA, Sandbrink HM, Fiers MW, Stiekema W, Lankhorst RM, Bron PA, Hoffer SM, Groot MN, Kerkhoven R, de Vries M, Ursing B, de Vos WM, Siezen RJ. (2003). Complete genome sequence of *Lactobacillus plantarum* WCFS1. *Proc Natl Acad Sci U S A*. 100(4):1990-1995.

Koenig RL, Ray JL, Maleki SJ, Smeltzer MS, Hurlburt BK (2004). *Staphylococcus aureus* AgrA binding to the RNAIII-*agr* regulatory region. *J Bacteriol*. 186(22):7549-7555.

Kornblum J, Kreiswirth BN, Projan SJ, Ross H Novick RP (1990). Agr: a polycistronic locus regulating exoprotein synthesis. Molecular Biology of the *Staphylococci* , pp. 373-402. Edited by R. P. Novick. New York:VCH Publishers.

Kosaka T, Nakayama S, Nakaya K, Yoshino S, Furukawa K (2007). Characterization of the *sol* operon in butanol-hyperproducing *Clostridium saccharoperbutylacetonicum* strain N1-4 and its degeneration mechanism. *Biosci Biotechnol Biochem*. 71(1):58-68..

Kroos L, Kunkel B, Losick R (1989). Switch protein alters specificity of RNA polymerase containing a compartment-specific sigma factor. *Science*. 243(4890):526-529.

Kunkel B, Losick R, Stragier P (1990). The *Bacillus subtilis* gene for the development transcription factor sigma K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene. *Genes Dev*. 4(4):525-535.

Lambowitz AM, Zimmerly S (2004). Mobile group II introns. *Annu Rev Genet*. 38:1-35.

Latifi A, Foglino M, Tanaka K, Williams P, Lazdunski A (1996). A hierarchical quorum-sensing cascade in *Pseudomonas aeruginosa* links the transcriptional activators LasR and RhIR (VsmR) to expression of the stationary-phase sigma factor RpoS. *Mol Microbiol*. 21(6):1137-1146.

Lauderdale KJ, Boles RB, Cheung AL, Horswill AR (2009). Interconnections between Sigma B, *agr*, and proteolytic activity in *Staphylococcus aureus* biofilm maturation. Infect Immun. 77(4):1623-1635.

Lazazzera BA, Solomon JM, Grossman AD (1997). An exported peptide functions intracellularly to contribute to cell density signaling in *B. subtilis*. *Cell*. 89(6):917-925.

Lazazzera BA, Kurtser IG, McQuade RS, Grossman AD. (1999). An autoregulatory circuit affecting peptide signaling in *Bacillus subtilis*. *J Bacteriol*. Sep;181(17):5193-5200.

Lazowska J, Meunier B, Macadre C (1994). Homing of a group II intron in yeast mitochondrial DNA is accompanied by unidirectional co-conversion of upstream-located markers. *EMBO J*. 13(20):4963-4972.

Lebeau C, Vandenesch F, Greenland T, Novick RP, Etienne J (1994). Coagulase expression in *Staphylococcus aureus* is positively and negatively modulated by an *agr*-dependent mechanism. *J Bacteriol*. 176(17):5534-5536.

Lee JH, Lequette Y, Greenberg EP (2006). Activity of purified QscR, a *Pseudomonas aeruginosa* orphan quorum-sensing transcription factor. *Mol Microbiol*. Jan;59(2):602-609.

Lee AS, Song KP (2005). LuxS/autoinducer-2 quorum sensing molecule regulates transcriptional virulence gene expression in *Clostridium difficile*. *Biochem Biophys Res Commun*. 335(3):659-666.

Lee, JY, Jang,YS, Lee, J, Papoutsakis, ET, Lee, SY (2009). Metabolic engineering of *Clostridium acetobutylicum* M5 for highly selective butanol production. *Biotechnol. J*. , 4,1432–1440.

Lenarcic R, Halbedel S, Visser L, Shaw M, Wu LJ, Errington J, Marenduzzo D, Hamoen LW (2009). Localisation of DivIVA by targeting to negatively curved membranes. *EMBO J*. 28(15):2272-2282.

Levin PA, Losick R (1996). Transcription factor Spo0A switches the localization of the cell division protein FtsZ from a medial to a bipolar pattern in *Bacillus subtilis*. *Genes Dev*. 10(4):478-488.

Levin PA, Losick R, Stragier P, Arigoni F (1997). Localization of the sporulation protein SpoIIE in *Bacillus subtilis* is dependent upon the cell division protein FtsZ. *Mol Microbiol*. 25(5):839-846.

Lina G, Jarraud S, Ji G, Greenland T, Pedraza A, Etienne J, Novick RP, Vandenesch F (1998). Transmembrane topology and histidine protein kinase activity of AgrC, the *agr* signal receptor in *Staphylococcus aureus*. *Mol Microbiol*. 28(3):655-662.

Liyanage H, Young M, Kashket ER (2000). Butanol tolerance of *Clostridium beijerinckii* NCIMB 8052 associated with down-regulation of *gldA* by antisense RNA. *J Mol Microbiol Biotechnol*. Jan;2(1):87-93.

Liyanage H, Kashket S, Young M, Kashket ER (2001). *Clostridium beijerinckii* and *Clostridium difficile* detoxify methylglyoxal by a novel mechanism involving glycerol dehydrogenase. *Appl Environ Microbiol*.67(5):2004-2010.

Londoño-Vallejo JA, Fréhel C, Stragier P (1997). SpoIIQ, a forespore-expressed gene required for engulfment in *Bacillus subtilis*. *Mol Microbiol*. 24(1):29-39.

Lopez-Diaz I, Clarke S, Mandelstam J (1986). *spoIID* operon of *Bacillus subtilis*: cloning and sequence. *J Gen Microbiol*. 132(2):341-354.

Lyon GJ, Muir TW (2003). Chemical signaling among bacteria and its inhibition. *Chem Biol*. 10(11):1007-1021.

Lyon GJ, Wright JS, Muir TW, Novick RP (2002). Key determinants of receptor activation in the *agr* autoinducing peptides of *Staphylococcus aureus*. *Biochemistry*. 41(31):10095-10104.

Lyristis M, Bryant AE, Sloan J, Awad MM, Nisbet IT, Stevens DL, Rood JI (1994). Identification and molecular analysis of a locus that regulates extracellular toxin production in *Clostridium perfringens*. *Mol Microbiol*. 12(5):761-777.

Maddox IS, Steiner E, Hirsch S, Wessner S, Gutierrez NA, Gapes JR, Schuster KC (2000). The cause of "acid-crash" and "acidogenic fermentations" during the batch acetone-butanol-ethanol (ABE-). fermentation process. J *Mol Microbiol* Biotechnol. 2(1):95-100.

Magnuson R, Solomon J, Grossman AD. (1994). Biochemical and genetic characterization of a competence pheromone from *B. subtilis*. *Cell*. 77(2):207-216.

Mandic-Mulec I, Doukhan L, Smith I (1995). The *Bacillus subtilis* SinR protein is a repressor of the key sporulation gene *spo0A*. *J Bacteriol* 177: 4619–4627.

Marquis KA, Burton BM, Nollmann M, Ptacin JL, Bustamante C, Ben-Yehuda S, Rudner DZ (2008). SpoIIIE strips proteins off the DNA during chromosome translocation. *Genes Dev*. 22(13):1786-1795.

Masuda ES, Anaguchi H, Sato T, Takeuchi M, Kobayashi Y (1990). Nucleotide sequence of the sporulation gene *spoIIGA* from *Bacillus subtilis*. *Nucleic Acids Res*. 18(3):657.

Matsushita C, Matsushita O, Koyama M, Okabe A (1994). A *Clostridium perfringens* vector for the selection of promoters. *Plasmid*. 31(3):317-319.

Mattson DM and Rogers P (1994). Analysis of Tn916-induced mutants of *Clostridium acetobutylicum* altered in solventogenesis and sporulation. *J Ind Microbiol*. 13:258-268

McNab R, Lamont RJ (2003). Microbial dinner-party conversations: the role of LuxS in interspecies communication. *J Med Microbiol*. 52(7):541-545.

Meisner J, Wang X, Serrano M, Henriques AO, Moran CP Jr (2008). A channel connecting the mother cell and forespore during bacterial endospore formation. *Proc Natl Acad Sci U S A*. 105(39):15100-15105.

Mermelstein LD, Papoutsakis ET (1993). In vivo methylation in *Escherichia coli* by the *Bacillus subtilis* phage phi 3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. *Appl Environ Microbiol*. 59(4):1077-1081.

Mermelstein LD, Papoutsakis ET, Petersen DJ, Bennett GN (1993). Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for increased solvent production by the enhancement of acetone formation enzyme activities using a synthetic acetone operon. *Biotechnol Bioeng*. 42:1053-1060.

Mermelstein LD, Welker NE, Bennett GN, Papoutsakis ET (1992). Expression of cloned homologous fermentative genes in *Clostridium acetobutylicum* ATCC 824. *Biotechnology* (N Y)..10(2):190-195.

Miller MB, Bassler BL (2001). Quorum sensing in bacteria. *Annu Rev Microbiol*. 55:165-199.

Min KT, Hilditch CM, Diederich B, Errington J, Yudkin MD(1993). Sigma F, the first compartment-specific transcription factor of *B.subtilis*, is regulated by an anti-sigma factor that is also a protein kinase. *Cell*. 74(4):735-742.

Moldover B, Piggot PJ, Yudkin MD (1991). Identification of the promoter and the transcriptional start site of the *spoVA* operon of *Bacillus subtilis* and *Bacillus licheniformis*. *J Gen Microbiol*. 137(3):527-531.

Moran JV, Zimmerly S, Eskes R, Kennell JC, Lambowitz AM, Butow RA, Perlman PS (1995). Mobile group II introns of yeast mitochondrial DNA are novel site-specific retroelements. *Mol Cell Biol*. 15(5):2828-2838.

Morfeldt E, Janzon L, Arvidson S, Löfdahl S (1988). Cloning of a chromosomal locus (*exp*). which regulates the expression of several exoprotein genes in *Staphylococcus aureus*. *Mol Gen Genet*. 211(3):435-440.

Morfeldt E, Panova-Sapundjieva I, Gustafsson B, Arvidson S (1996a). Detection of the response regulator AgrA in the cytosolic fraction of *Staphylococcus aureus* by monoclonal antibodies. *FEMS Microbiol Lett*. 143(2-3):195-201.

Morfeldt E, Tegmark K, Arvidson S (1996b). Transcriptional control of the *agr*-dependent virulence gene regulator, RNAIII, in *Staphylococcus aureus*. *Mol Microbiol*. 21(6):1227-1237.

Moriyama R, Fukuoka H, Miyata S, Kudoh S, Hattori A, Kozuka S, Yasuda Y, Tochikubo K, Makino S (1999). Expression of a germination-specific amidase, SleB, of Bacilli in the forespore compartment of sporulating cells and its localization on the exterior side of the cortex in dormant spores. *J Bacteriol*. 181(8):2373-2378.

Mullany P, Wilks M, Tabaqchali S (1991). Transfer of Tn*916* and Tn*916* delta E into *Clostridium difficile*: demonstration of a hot-spot for these elements in the *C. difficile* genome*. FEMS Microbiol Lett*. 63(2-3):191-194.

Narberhaus F, Giebeler K, Bahl H (1992). Molecular characterization of the *dnaK* gene region of *Clostridium acetobutylicum*, including *grpE*, *dnaJ*, and a new heat shock gene. *J Bacteriol*. 174(10):3290-3299.

Nair RV, Green EM, Watson DE, Bennett GN, Papoutsakis ET (1999). Regulation of the *sol* locus genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 by a putative transcriptional repressor. *J Bacteriol*. 181(1):319-330.

Nakayama J, Cao Y, Horii T, Sakuda S, Nagasawa H. (2001). Chemical synthesis and biological activity of the gelatinase biosynthesis-activating pheromone of *Enterococcus faecalis* and its analogs. *Biosci Biotechnol Biochem*. 65(10):2322-2325.

Nakayama S, Kosaka T, Hirakawa H, Matsuura K, Yoshino S, Furukawa K (2008). Metabolic engineering for solvent productivity by downregulation of the hydrogenase gene cluster *hupCBA* in *Clostridium saccharoperbutylacetonicum* strain N1-4. *Appl Microbiol Biotechnol*. 78(3):483-493.

Nealson KH, Platt T, Hastings JW (1970). Cellular control of the synthesis and activity of the bacterial luminescent system. *J Bacteriol*. Oct;104(1):313-322.

Nes IF, Diep DB, Håvarstein LS, Brurberg MB, Eijsink V, Holo H. (1996). Biosynthesis of bacteriocins in lactic acid bacteria. *Antonie Van Leeuwenhoek* 70, 113–128.

Nölling J, Breton G, Omelchenko MV, Makarova KS, Zeng Q, Gibson R, Lee HM, Dubois J, Qiu D, Hitti J, Wolf YI, Tatusov RL, Sabathe F, Doucette-Stamm L, Soucaille P, Daly MJ, Bennett GN, Koonin EV, Smith DR (2001). Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J Bacteriol*. 183(16):4823-4838.

Novick RP, Projan SJ, Kornblum J, Ross HF, Ji G, Kreiswirth B, Vandenesch F, Moghazeh S (1995). The *agr* P2 operon: an autocatalytic sensory transduction system in *Staphylococcus aureus*. *Mol Gen Genet*. 248(4):446-458.

O'Brien DK, Melville SB (2004). Effects of *Clostridium perfringens* alpha-toxin (PLC). and perfringolysin O (PFO). on cytotoxicity to macrophages, on escape from the phagosomes of macrophages, and on persistence of *C. perfringens* in host tissues. Infect Immun. 72(9):5204-5215.

O'Connor JR, Lyras D, Farrow KA, Adams V, Powell DR, Hinds J, Cheung JK, Rood JI (2006). Construction and analysis of chromosomal *Clostridium difficile* mutants. *Mol Microbiol*. 61(5):1335-1351.

Ohtani K, Hayashi H, Shimizu T (2002). The *luxS* gene is involved in cell-cell signalling for toxin production in *Clostridium perfringens*. *Mol Microbiol*. 44(1):171-179.

Ohtani K, Yuan Y, Hassan S, Wang R, Wang Y, Shimizu T (2009). Virulence gene regulation by the *agr* system in *Clostridium perfringens*. *J Bacteriol*. 191(12):3919-3927.

Oinuma K, Greenberg EP (2011). Acyl-homoserine lactone binding to and stability of the orphan *Pseudomonas aeruginosa* quorum-sensing signal receptor QscR. *J Bacteriol*. 193(2):421-8.

Oultram JD, Loughlin M, Swinfield TJ, Brehm JK, Thompson DE, and Minton NP (1988a). Introduction of plasmids into whole cells of *Clostridium acetobutylicum* by electroporation. *FEMS Microbiol Lett*. 56:83-88.

Oultram JD, Peck H, Brehm JK, Thompson DE, Swinfield TJ, Minton NP (1988b). Introduction of genes for leucine biosynthesis from *Clostridium pasteurianum* into *C. acetobutylicum* by cointegrate conjugal transfer. Mol Gen Genet. 214(1):177-179.

Paredes CJ, Rigoutsos I, Papoutsakis ET (2004). Transcriptional organization of the *Clostridium acetobutylicum* genome. *Nucleic Acids Res*. 32(6):1973-1981.

Paredes CJ, Senger RS, Spath IS, Borden JR, Sillers R, Papoutsakis ET (2007). A general framework for designing and validating oligomer-based DNA microarrays and its application to *Clostridium acetobutylicum*. *Appl Environ Microbiol*. 73(14):4631-4638.

Partridge SR, Errington J (1993). The importance of morphological events and intercellular interactions in the regulation of prespore-specific gene expression during sporulation in *Bacillus subtilis*. *Mol Microbiol*. 8(5):945-955.

Peng HL, Novick RP, Kreiswirth B, Kornblum J, Schlievert P (1988). Cloning, characterization, and sequencing of an accessory gene regulator (*agr*). in *Staphylococcus aureus*. *J Bacteriol*. 170(9):4365-4372.

Perret S, Maamar H, Bélaich JP, Tardif C (2004). Use of antisense RNA to modify the composition of cellulosomes produced by *Clostridium cellulolyticum*. *Mol Microbiol*. 51(2):599-607.

Perutka J, Wang W, Goerlitz D, Lambowitz AM (2004). Use of computer-designed group II introns to disrupt *Escherichia coli* DExH/D-box protein and DNA helicase genes. *J Mol Biol*. 336(2):421-439.**Pesci EC, Pearson JP, Seed PC, Iglewski BH (1997)** Regulation of *las* and *rhl* quorum sensing in *Pseudomonas aeruginosa*. *J Bacteriol*. 179(10):3127-132.

Pesci EC, Milbank JB, Pearson JP, McKnight S, Kende AS, Greenberg EP, Iglewski BH (1999). Quinolone signaling in the cell-to-cell communication system of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci.U. S. A.* 96:11229–11234.

Petersen D J, Cary JW, Vanderleyden J, Bennett GN (1993). Sequence and arrangement of genes encoding enzymes of the acetone-production pathway of *Clostridium acetobutylicum* ATCC 824.*Gene* 123: 93-97.

Piggot PJ, Coote JG (1976). Genetic aspects of bacterial endospore formation. *Bacteriol Rev*. 40(4):908-962..

Pillai SK, Sakoulas G, Eliopoulos GM, Moellering RC Jr, Murray BE, Inouye RT. (2004). Effects of glucose on *fsr*-mediated biofilm formation in *Enterococcus faecalis*. *J Infect Dis* 190, 967–970.

Pogliano J, Osborne N, Sharp MD, Abanes-De Mello A, Perez A, Sun YL, Pogliano K (1999). A vital stain for studying membrane dynamics in bacteria: a novel mechanism controlling septation during *Bacillus subtilis* sporulation. *Mol Microbiol*. 31(4):1149-1159.

Purdy D, O'Keeffe TA, Elmore M, Herbert M, McLeod A, Bokori-Brown M, Ostrowski A, Minton NP (2002). Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier. *Mol Microbiol*. 46(2):439-452.

Qin X, Singh KV, Weinstock GM, Murray BE (2000). Effects of *Enterococcus faecalis fsr* genes on production of gelatinase and a serine protease and virulence. *Infect Immun* 68, 2579–2586.

Qin X, Singh KV, Weinstock GM, Murray BE. (2001). Characterization of *fsr*, a regulator controlling expression of gelatinase and serine protease in *Enterococcus faecalis* OG1RF. *J Bacteriol* 183, 3372–3382.

Qiu R, Pei W, Zhang L, Lin J, Ji G (2005). Identification of the putative staphylococcal AgrB catalytic residues involving the proteolytic cleavage of AgrD to generate autoinducing peptide. J Biol Chem. Apr 29;280(17):16695-16704.

Queck SY, Jameson-Lee M, Villaruz AE, Bach TH, Khan BA, Sturdevant DE, Ricklefs SM, Li M, Otto M (2008). RNAIII-independent target gene control by the *agr* quorum-sensing system: insight into the evolution of virulence regulation in *Staphylococcus aureus*. *Mol Cell.*32(1):150-158.

Raju D, Setlow P, Sarker MR (2007). Antisense-RNA-mediated decreased synthesis of small, acid-soluble spore proteins leads to decreased resistance of *Clostridium perfringens* spores to moist heat and UV radiation. *Appl Environ Microbiol*. 73(7):2048-2053.

Raju D, Waters M, Setlow P, Sarker MR (2006). Investigating the role of small, acid-soluble spore proteins (SASPs). in the resistance of *Clostridium perfringens* spores to heat. *BMC Microbiol*. 6:50.

Rangarajan ES, Asinas A, Proteau A, Munger C, Baardsnes J, Iannuzzi P, Matte A, Cygler M (2008). Structure of [NiFe] hydrogenase maturation protein HypE from *Escherichia coli* and its interaction with HypF. *J Bacteriol*. 190(4):1447-58.

Rasmussen LC, Sperling-Petersen HU, Mortensen KK (2007). Hitting bacteria at the heart of the central dogma: sequence-specific inhibition. *Microb Cell Fact*. (10)6-24.

Read TD, Thomas AT, Wilkins BM. (1992). Evasion of type I and type II DNA restriction systems by IncI1 plasmid CoIIb-P9 during transfer by bacterial conjugation. *Mol Microbiol*. 6(14):1933-1941.

Recsei P, Kreiswirth B, O'Reilly M, Schlievert P, Gruss A, Novick RP(1986). Regulation of exoprotein gene expression in *Staphylococcus aureus* by agr. *Mol Gen Genet*. 202(1):58-61.

Reid SJ, Rafudeen MS, Leat NG (1999). The genes controlling sucrose utilization in *Clostridium beijerinckii* NCIMB 8052 constitute an operon. *Microbiology*. 145 (6):1461-1472.

Reyes D, Andrey DO, Monod A, Kelley WL, Zhang G, Cheung AL (2011). Coordinated regulation by AgrA, SarA, and SarR to control agr expression in *Staphylococcus aureus*. *J Bacteriol*. 193(21):6020-6031.

Riedel CU, Monk IR, Casey PG, Waidmann MS, Gahan CG, Hill C (2009). *AgrD*-dependent quorum sensing affects biofilm formation, invasion, virulence and global gene expression profiles in *Listeria monocytogenes*. *Mol Microbiol*. 71(5):1177-1189.

Rieu A, Weidmann S, Garmyn D, Piveteau P, Guzzo J (2007). Agr system of *Listeria monocytogenes* EGD-e: role in adherence and differential expression pattern. *Appl Environ Microbiol*. 73(19):6125-6133.

Rieu A, Briandet R, Habimana O, Garmyn D, Guzzo J, Piveteau P (2008). *Listeria monocytogenes* EGD-e biofilms: no mushrooms but a network of knitted chains. *Appl Environ Microbiol*. Jul;74(14):4491-4497.

Roberts AP, Hennequin C, Elmore M, Collignon A, Karjalainen T, Minton N, Mullany P (2003). Development of an integrative vector for the expression of antisense RNA in *Clostridium difficile*. *J Microbiol Methods*. 55(3):617-624.

Robertson JB, Gocht M, Marahiel MA, Zuber P (1989). AbrB, a regulator of gene expression in *Bacillus*, interacts with the transcription initiation regions of a sporulation gene and an antibiotic biosynthesis gene. *Proc Natl Acad Sci U S A*. 86(21):8457-8461.

Rong S, Rosenkrantz MS, Sonenshein AL (1986). Transcriptional control of the *Bacillus subtilis spoIID* gene. *J Bacteriol*. 165(3):771-779.

Rood JI, Cole ST (1991). Molecular genetics and pathogenesis of *Clostridium perfringens*. *Microbiol Rev*. 55(4):621-648. Review.

Ryter A (1965). Morphologic study of the sporulation of *bacillus subtilis*. *Ann Inst Pasteur* (Paris). 108:40-60.

Saint-Prix, F (2006). Development of a system for chromosomal Integration and gene deletion in *Clostridium acetobutylicum* ATCC 824. In:Abstracts of the 9th international workshop on the regulation of metabolism, genetics and development of the solvent- and acid-producing Clostridia. Houston:USA

Sambrook, J., Fritsch, E. F. & Maniatis, T (1989). Molecular Cloning: a Laboratory Manual, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sarker MR, Carman RJ, McClane BA (1999). Inactivation of the gene (*cpe*). encoding *Clostridium perfringens* enterotoxin eliminates the ability of two *cpe*-positive *C. perfringens* type A human gastrointestinal disease isolates to affect rabbit ileal loops. *Mol Microbiol*. 33(5):946-58. Erratum in: *Mol Microbiol* 35(1):249.

Sarker R I, Ogawa W, Tsuda M, Tanaka S,Tsuchiya T (1994). Characterization of a glucose transport system in *Vibrio parahaemolyticus*. *J Bacteriol*. 176(23): 7378-7382

Scotcher MC, Bennett GN (2005). SpoIIE regulates sporulation but does not directly affect solventogenesis in *Clostridium acetobutylicum* ATCC 824. *J Bacteriol*. 187(6):1930-1936.

Scotcher MC, Bennett GN (2008). Activity of *abrB310* promoter in wild type and spo0A-deficient strains of *Clostridium acetobutylicum*. *J Ind Microbiol Biotechnol*. 35(7):743-750.

Scotcher MC, Rudolph FB, Bennett GN (2005). Expression of *abrB310* and *sinR*, and effects of decreased *abrB310* expression on the transition from acidogenesis to solventogenesis, in *Clostridium acetobutylicum* ATCC 824. *Appl Environ Microb* 71(4):1987-1995.

Sebaihia M, Wren BW, Mullany P, Fairweather NF, Minton N, Stabler R, Thomson NR, Roberts AP, Cerdeño-Tárraga AM, Wang H, Holden MT, Wright A, Churcher C, Quail MA, Baker S, Bason N, Brooks K, Chillingworth T, Cronin A, Davis P, Dowd L, Fraser A, Feltwell T, Hance Z, Holroyd S, Jagels K, Moule S, Mungall K, Price C, Rabbinowitsch E, Sharp S, Simmonds M, Stevens K, Unwin L, Whithead S, Dupuy B, Dougan G, Barrell B, Parkhill J (2005). The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nat Genet*. 38(7):779-786.

Serrano M, Côrte L, Opdyke J, Moran CP Jr, Henriques AO (2003). Expression of *spoIIIJ* in the prespore is sufficient for activation of sigma G and for sporulation in *Bacillus subtilis*. *J Bacteriol*. 185(13):3905-3917.

Serrano M, Neves A, Soares CM, Moran CP Jr, Henriques AO (2004). Role of the anti-sigma factor SpoIIAB in regulation of sigmaG during *Bacillus subtilis* sporulation. *J Bacteriol*. 186(12):4000-4013.

Serrano M, Vieira F, Moran CP Jr, Henriques AO (2008). Processing of a membrane protein required for cell-to-cell signaling during endospore formation in *Bacillus subtilis*. *J Bacteriol*.190(23):7786-7796.

Sharp MD, Pogliano K (1999). An in vivo membrane fusion assay implicates SpoIIIE in the final stages of engulfment during *Bacillus subtilis* sporulation*. Proc Natl Acad Sci U S A*. 96(25):14553-14558.

Sharp MD, Pogliano K (2003). The membrane domain of SpoIIIE is required for membrane fusion during *Bacillus subtilis* sporulation. *J Bacteriol*. 185(6):2005-2008.

Shi Z, Blaschek HP. (2008). Transcriptional analysis of *Clostridium beijerinckii* NCIMB 8052 and the hyper-butanol-producing mutant BA101 during the shift from acidogenesis to solventogenesis. *Appl Environ Microbiol*. 74(24):7709-7714.

Shimizu T, Ba-Thein W, Tamaki M, Hayashi H (1994). The *virR* gene, a member of a class of two-component response regulators, regulates the production of perfringolysin O, collagenase, and hemagglutinin in *Clostridium perfringens*. *J Bacteriol*. 176(6):1616-1623.

Shimizu T, Yaguchi H, Ohtani K, Banu S, Hayashi H (2002). Clostridial VirR/VirS regulon involves a regulatory RNA molecule for expression of toxins. *Mol Microbiol*. 43(1):257-265.

Sikkema J, de Bont JA, Poolman B (1995). Mechanisms of membrane toxicity of hydrocarbons. *Microbiol Rev*. 59(2):201-222.

Sillers R, Al-Hinai MA, Papoutsakis ET. (2009). Aldehyde-alcohol dehydrogenase and/or thiolase overexpression coupled with CoA transferase downregulation lead to higher alcohol titers and selectivity in *Clostridium acetobutylicum* fermentations. *Biotechnol Bioeng*. 102(1):38-49.

Smith K, Bayer ME, Youngman P (1993). Physical and functional characterization of the *Bacillus subtilis spoIIM* gene. *J Bacteriol*. 175(11):3607-3617.

Smith K, Youngman P (1993). Evidence that the *spoIIM* gene of *Bacillus subtilis* is transcribed by RNA polymerase associated with sigma E. *J Bacteriol*. 175(11):3618-3627.

Solomon JM, Magnuson R, Srivastava A, Grossman AD (1995). Convergent sensing pathways mediate response to two extracellular competence factors in *Bacillus subtilis*. *Genes Dev*. 9(5):547-558.

Steil L, Serrano M, Henriques AO, Völker U (2005). Genome-wide analysis of temporally regulated and compartment-specific gene expression in sporulating cells of *Bacillus subtilis*. *Microbiology*. 151(2):399-420.

Steiner E, Scott J, Minton NP, Winzer K (2012). An *agr* Quorum Sensing System That Regulates Granulose Formation and Sporulation in *Clostridium acetobutylicum*. *Appl Environ Microbiol*. 78(4):1113-22.

Stephenson K, Hoch JA (2002). Evolution of signalling in the sporulation phosphorelay. *Mol Microbiol*. 46(2):297-304.

Stragier P, Bouvier J, Bonamy C, Szulmajster J. (1984). A developmental gene product of *Bacillus subtilis* homologous to the sigma factor of *Escherichia coli*. *Nature*. 312(5992):376-378.

Stragier P, Kunkel B, Kroos L, Losick R (1989). Chromosomal rearrangement generating a composite gene for a developmental transcription factor. *Science*. 243(4890):507-512.

Stragier P, Losick R (1996). Molecular genetics of sporulation in *Bacillus subtilis*. *Annu Rev Genet*. 30:297-241.

Sturme MH, Kleerebezem M, Nakayama J, Akkermans AD, Vaugha EE, de Vos WM. Antonie Van Leeuwenhoek. (2002). Cell to cell communication by autoinducing peptides in gram-positive bacteria. *Antonie Van Leeuwenhoek*. 81(1-4):233-243.

Sturme MH, Francke C, Siezen RJ, de Vos WM, Kleerebezem M. (2007). Making sense of quorum sensing in *lactobacilli*: a special focus on *Lactobacillus plantarum* WCFS1. *Microbiology*. 153(Pt 12):3939-3947.

Sullivan L, Paredes C, Papoutsakis ET, Bennett G (2007) Analysis of the clostridial hydrophobic with a conserved tryptophan family (ChW) of proteins in *Clostridium acetobutylicum* with emphasis on ChW14 and ChW16/17. *Enzyme and Microbial Technology* 42 29–43

Sun DX, Cabrera-Martinez RM, Setlow P (1991). Control of transcription of the *Bacillus subtilis spoIIIG* gene, which codes for the forespore-specific transcription factor sigma G. *J Bacteriol*. 173(9):2977-2984.

Sun DX, Stragier P, Setlow P (1989). Identification of a new sigma-factor involved in compartmentalized gene expression during sporulation of *Bacillus subtilis*. *Genes Dev*. 3(2):141-149.

Szurmant H, Ordal GW (2004). Diversity in chemotaxis mechanisms among the bacteria and archaea. *Microbiol Mol Biol Rev*. 68(2):301-319.

Takamatsu H, Chikahiro Y, Kodama T, Koide H, Kozuka S, Tochikubo K, Watabe K (1998). A spore coat protein, CotS, of *Bacillus subtilis* is synthesized under the regulation of sigmaK and GerE during development and is located in the inner coat layer of spores. *J Bacteriol*. 180(11):2968-2974.

Takano E, Nihira T, Hara Y, Jones JJ, Gershater CJ, Yamada Y, Bibb M (2000). Purification and structural determination of SCB1, a gamma-butyrolactone that elicits antibiotic production in *Streptomyces coelicolor* A3(2). J Biol Chem. 275(15):11010-11016.

Tatsuki T, Imagawa T, Kitajima H, Higashi Y, Chazono M, Yanagase Y, Amano T (1981). Further studies on complementation between mutants of *Clostridium perfringens*. *Biken J*. 24(1-2):1-11.

Thoendel M, Kavanaugh JS, Flack CE, Horswill AR (2011). Peptide signaling in the *staphylococci*. Chem Rev. 111(1):117-151.

Thomaides HB, Freeman M, El Karoui M, Errington J (2001). Division site selection protein DivIVA of *Bacillus subtilis* has a second distinct function in chromosome segregation during sporulation. *Genes Dev*. 15(13):1662-73.

Thormann K and Dürre P. (2001). Orf5/SolR: a transcriptional repressor of the *sol* operon of *Clostridium acetobutylicum*? *J Ind Microbiol. Biotechnol*., 27: 307–313.

Thormann K, Feustel L, Lorenz K, Nakotte S, Dürre P. (2002). Control of butanol formation in *Clostridium acetobutylicum* by transcriptional activation. *J Bacteriol*. 184(7):1966-1973.

Tomas CA, Welker NE, Papoutsakis ET (2003a). Overexpression of *groESL* in *Clostridium acetobutylicum* results in increased solvent production and tolerance, prolonged metabolism, and changes in the cell's transcriptional program. *Appl Environ Microbiol*. 69(8):4951-4965.

Tomas CA, Alsaker KV, Bonarius HP, Hendriksen WT, Yang H, Beamish JA, Paredes CJ, Papoutsakis ET (2003b). DNA array-based transcriptional analysis of asporogenous, nonsolventogenic *Clostridium acetobutylicum* strains SKO1 and M5. *J Bacteriol*. 185(15):4539-4547.

Tortosa P, Logsdon L, Kraigher B, Itoh Y, Mandic-Mulec I, Dubnau D. (2001). Specificity and genetic polymorphism of the *Bacillus* competence quorum-sensing system. *J Bacteriol*. 183(2):451-460.

Tovar-Rojo F, Chander M, Setlow B, Setlow P (2002). The products of the *spoVA* operon are involved in dipicolinic acid uptake into developing spores of *Bacillus subtilis*. *J Bacteriol*. 184(2):584-587.

Trempy JE, Haldenwang WG (1985). Sigma 29-like protein is a common sporulation-specific element in bacteria of the genus *Bacillus*. *J Bacteriol*. 164(3):1356-1358.

Tummala SB, Welker NE, Papoutsakis ET (1999). Development and characterization of a gene expression reporter system for *Clostridium acetobutylicum* ATCC 824. *Appl Environ Microbiol*. 65(9):3793-3799.

Tummala SB, Junne SG, Paredes CJ, Papoutsakis ET (2003a). Transcriptional analysis of product-concentration driven changes in cellular programs of recombinant *Clostridium acetobutylicum* strains. *Biotechnol Bioeng*. 84(7):842-854.

Tummala SB, Junne SG, Papoutsakis ET (2003b). Antisense RNA downregulation of coenzyme A transferase combined with alcohol-aldehyde dehydrogenase overexpression leads to predominantly alcohologenic *Clostridium acetobutylicum* fermentations. *J Bacteriol*. 185(12):3644-3653.

Tummala SB, Welker NE, Papoutsakis ET (2003c) Design of antisense RNA constructs for downregulation of the acetone formation pathway of *Clostridium acetobutylicum*. *J Bacteriol*. 185(6):1923-1934.

Tyurin M, Padda R, Huang KX, Wardwell S, Caprette D, Bennett GN (2000). Electrotransformation of *Clostridium acetobutylicum* ATCC 824 using high-voltage radio frequency modulated square pulses. *J Appl Microbiol*. 88(2):220-227.

Underwood S, Guan S, Vijayasubhash V, Baines SD, Graham L, Lewis RJ, Wilcox MH, Stephenson K (2009). Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. *J Bacteriol*. 191(23):7296-7305

Vandenesch F, Kornblum J, Novick RP (1991). A temporal signal, independent of *agr*, is required for *hla* but not *spa* transcription in *Staphylococcus aureus*. *J Bacteriol*. 173(20):6313-6320.

Varga JJ, Nguyen V, O'Brien DK, Rodgers K, Walker RA, Melville SB (2006). Type IV pili-dependent gliding motility in the Gram-positive pathogen *Clostridium perfringens* and other *Clostridia*. *Mol Microbiol*. 62(3):680-694.

Vasudevan P, Weaver A, Reichert ED, Linnstaedt SD, Popham DL (2007). Spore cortex formation in *Bacillus subtilis* is regulated by accumulation of peptidoglycan precursors under the control of sigma K. *Mol Microbiol*. 65(6):1582-1594.

Vidal JE, Chen J, Li J, McClane BA (2009). Use of an EZ-Tn5-based random mutagenesis system to identify a novel toxin regulatory locus in *Clostridium perfringens* strain 13. *PLoS One*. 4(7):e6232.

von Bodman SB, Willey JM, Diggle SP (2008). Cell-cell communication in bacteria: united we stand. *J Bacteriol*. 190(13):4377-4391**Walter KA,**

Mermelstein LD, Papoutsakis ET (1994). Studies of recombinant *Clostridium acetobutylicum* with increased dosages of butyrate formation genes. *Ann N Y Acad Sci*. 721:69-72.

Wang H, Roberts AP, Mullany P (2000). DNA sequence of the insertional hot spot of Tn*916* in the *Clostridium difficile* genome and discovery of a Tn*916*-like element in an environmental isolate integrated in the same hot spot. *FEMS Microbiol Lett*. 192(1):15-20.

Wang L, Grau R, Perego M, Hoch JA (1997). A novel histidine kinase inhibitor regulating development in *Bacillus subtilis*. *Genes Dev*. 11(19):2569-2579.

Wang L, Perpich J, Driks A, Kroos L (2007). One perturbation of the mother cell gene regulatory network suppresses the effects of another during sporulation of *Bacillus subtilis*. *J Bacteriol*. 189(23):8467-84 73.

Wang X, Lutkenhaus J (1993). The FtsZ protein of *Bacillus subtilis* is localized at the division site and has GTPase activity that is dependent upon FtsZ concentration. *Mol Microbiol*. 9(3):435-4342.

Wanner BL (1992). Is cross regulation by phosphorylation of two-component response regulator proteins important in bacteria? *J Bacteriol*. 174(7):2053-2058.

Welch M, Oosawa K, Aizawa SI, Eisenbach M (1994). Effects of phosphorylation, Mg^{2+}, and conformation of the chemotaxis protein CheY on its binding to the flagellar switch protein FliM. *Biochemistry*. 33(34):10470-10476.

Wilkins BM (2002). Plasmid promiscuity: meeting the challenge of DNA immigration control. *Environ Microbiol*, 4:495-500.

Wilkinson S,Young M (1994). Targeted integration of genes into the *Clostridium acetobutylicum* chromosome. *Microbiology* 140 89-95.

Williams P (2007a). Quorum sensing, communication and cross-kingdom signalling in the bacterial world. *Microbiology*. 153:3923-38.

Williams P, Winzer K, Chan WC, Cámara M. (2007b). Look who's talking: communication and quorum sensing in the bacterial world. *Philos Trans R Soc Lond B Biol Sci*. 362(1483):1119-1134.

Williams P, Camara M, Hardman A, Swift S, Milton D, Hope VJ, Winzer K, Middleton B, Pritchard DI, Bycroft BW (2000). Quorum sensing and the population-dependent control of virulence. *Philos Trans R Soc Lond B Biol Sci*. 355(1397):667-680.

Williams T, Bauer S, Beier D, Kuhn M. (2005). Construction and characterization of *Listeria monocytogenes* mutants with in-frame deletions in the response regulator genes identified in the genome sequence. *Infect Immun*. 73: 3152–3159.

Winzer K, Hardie KR, Williams P (2002). Bacterial cell-to-cell communication: sorry, can't talk now - gone to lunch! *Curr Opin Microbiol*. 5(2):216-222.

Wörner K, Szurmant H, Chiang C, Hoch JA (2006). Phosphorylation and functional analysis of the sporulation initiation factor Spo0A from *Clostridium botulinum*. *Mol Microbiol*. 59(3):1000-1012.

Wright JS 3rd, Traber KE, Corrigan R, Benson SA, Musser JM, Novick RP (2005). The *agr* radiation: an early event in the evolution of *staphylococci*. *J Bacteriol*. 187(16):5585-5594

Wu LJ, Errington J (1998). Use of asymmetric cell division and *spoIIIE* mutants to probe chromosome orientation and organization in *Bacillus subtilis*. *Mol Microbiol*. 27(4):777-786.

Wu LJ, Errington J (2003). RacA and the Soj-Spo0J system combine to effect polar chromosome segregation in sporulating *Bacillus subtilis*. *Mol Microbiol*. 49(6):1463-1475.

Wuster A, Babu MM (2008). Conservation and evolutionary dynamics of the *agr* cell-to-cell communication system across firmicutes. *J Bacteriol*. 190(2):743-746.

Xiao G, Deziel E, He J, Lepine F, Lesic B, Castonguay MH, Milot S, Tampakaki AP, Stachel SE, Rahme LG (2006). MvfR, a key *Pseudomonas aeruginosa* pathogenicity LTTR-class regulatory protein, has dual ligands. *Mol. Microbiol*. 62:1689–1699.

Yao J, Zhong J, Fang Y, Geisinger E, Novick RP, Lambowitz AM (2006). Use of targetrons to disrupt essential and nonessential genes in *Staphylococcus aureus* reveals temperature sensitivity of Ll.LtrB group II intron splicing. *RNA*. 12(7):1271-1281.

Yarwood JM, Bartels DJ, Volper EM, Greenberg EP (2004). Quorum sensing in *Staphylococcus aureus* biofilms. *J Bacteriol*. 186(6):1838-1850.

Yarwood JM, McCormick JK, Paustian ML, Kapur V, Schlievert PM (2002). Repression of the *Staphylococcus aureus* accessory gene regulator in serum and in vivo. *J Bacteriol*. 184(4):1095-1101.

Yarwood JM, Schlievert PM (2003). Quorum sensing in *Staphylococcus* infections. *J Clin Invest*. 112(11):1620-1625.

Zanen G, Antelmann H, Westers H, Hecker M, van Dijl JM, Quax WJ (2004). FlhF, the third signal recognition particle-GTPase of *Bacillus subtilis*, is dispensable for protein secretion. *J Bacteriol*. 186(17):5956-5960.

Zhang L, Gray L, Novick RP, Ji G (2002). Transmembrane topology of AgrB, the protein involved in the post-translational modification of AgrD in *Staphylococcus aureus*. *J Biol Chem*. Sep 20;277(38):34736-34742.

Zhang L, Lin J, Ji G (2004). Membrane anchoring of the AgrD N-terminal amphipathic region is required for its processing to produce a quorum-sensing pheromone in *Staphylococcus aureus*. *J Biol Chem*. 279(19):19448-19456.

Zhao Y, Hindorff LA, Chuang A, Monroe-Augustus M, Lyristis M, Harrison ML, Rudolph FB, Bennett GN (2003). Expression of a cloned cyclopropane fatty acid synthase gene reduces solvent formation in *Clostridium acetobutylicum* ATCC 824. *Appl Environ Microbiol*. 69(5):2831-2841.

Zhao L, Montville TJ, Schaffner DW. (2006). Evidence for quorum sensing in *Clostridium botulinum* 56A*. Lett. Appl. Microbiol*.;42:54–58.

Zheng L, Halberg R, Roels S, Ichikawa H, Kroos L, Losick R (1992). Sporulation regulatory protein GerE from *Bacillus subtilis* binds to and can activate or repress transcription from promoters for mother-cell-specific genes. *J Mol Biol*. 226(4):1037-1050.

Zhong J, Karberg M, Lambowitz AM (2003). Targeted and random bacterial gene disruption using a group II intron (targetron). vector containing a retrotransposition-activated selectable marker. *Nucleic Acids Res*. 31(6):1656-1664.

Zuberi AR, Ying C, Bischoff DS, Ordal GW (1991). Gene-protein relationships in the flagellar hook-basal body complex of *Bacillus subtilis*: sequences of the *flgB*, *flgC*, *flgG*, *fliE* and *fliF* genes. *Gene*. 101(1):23-31.

Appendix I – z scores of the three *agr* mutants with the RAM for the levels of product formation

Ethanol

	agrA z score	agrA + Comp vector z score	agrA + plasmid control z score	agrB z score	agrB + Comp vector z score	agrB + plasmid control z score	agrC z score	agrC + Comp vector z score	agrC + plasmid control z score
2h	1.964	0.655	1.528	1.964	0.655	1.964	1.964	1.964	1.964
4h	0.218	1.964	1.964	1.964	0.655	1.964	1.964	1.964	1.964
6h	1.964	1.964	1.964	1.964	0.218	1.964	1.091	1.964	1.964
72h	0.655	1.528	0.218	0.218	1.964	1.964	1.964	1.964	1.964

Acetone

	agrA z score	agrA + Comp vector z score	agrA + plasmid control z score	agrB z score	agrB + Comp vector z score	agrB + plasmid control z score	agrC z score	agrC + Comp vector z score	agrC + plasmid control z score
2h	1.528	1.964	0.655	1.964	1.964	0.655	1.964	0.655	0.655
4h	1.964	1.964	1.964	1.964	1.964	1.964	1.964	1.964	0.655
6h	1.964	1.964	0.218	1.964	0.655	0.655	1.964	1.964	1.964
72h	1.964	1.964	1.964	1.964	1.964	0.655	1.964	1.964	1.964

Acetate

	agrA z score	agrA + Comp vector z score	agrA + plasmid control z score	agrB z score	agrB + Comp vector z score	agrB + plasmid control z score	agrC z score	agrC + Comp vector z score	agrC + plasmid control z score
2h	1.528	0.655	1.091	0.655	0.218	1.964	0.655	1.964	0.218
4h	0.655	0.218	0.655	1.964	1.091	1.528	0.218	1.964	1.964
6h	0.655	1.528	1.964	1.964	0.218	1.964	0.655	1.964	1.528
72h	1.964	1.528	1.528	0.655	1.964	1.091	0.655	0.655	1.091

Butanol

	agrA z score	agrA + Comp vector z score	agrA + plasmid control z score	agrB z score	agrB + Comp vector z score	agrB + plasmid control z score	agrC z score	agrC + Comp vector z score	agrC + plasmid control z score
2h	1.964	1.964	1.964	1.964	0.655	0.655	1.964	1.964	1.964
4h	1.964	1.964	0.218	1.964	1.964	1.091	1.091	1.964	1.964
6h	1.964	1.964	1.964	1.964	1.964	1.964	0.655	1.964	1.091
72h	1.091	1.528	1.964	1.528	1.091	1.964	1.091	0.655	1.964

Butyrate

	agrA z score	agrA + Comp vector z score	agrA + plasmid control z score	agrB z score	agrB + Comp vector z score	agrB + plasmid control z score	agrC z score	agrC + Comp vector z score	agrC + plasmid control z score
2h	1.528	1.964	1.964	1.964	1.091	1.964	1.091	1.964	1.964
4h	1.964	1.964	1.964	0.655	1.964	1.528	1.091	1.964	1.964
6h	1.964	1.964	1.964	1.964	1.091	1.964	0.655	1.964	1.964
72h	1.964	1.528	0.218	0.218	1.964	1.528	0.655	0.218	1.528

Appendix II – z scores of the four agr Mutants containing no selective marker for the levels of product formation

Ethanol

	AgrA z score	AgrB z Score	AgrC z score	AgrA deletion mutant z score
2h	1.9640	1.9640	1.9640	1.9640
4h	1.9640	1.9640	1.9640	1.9640
6h	1.0911	1.9640	1.0911	1.0911
72h	1.9640	1.5275	1.9640	0.2182

Acetone

	AgrA z score	AgrB z Score	AgrC z score	AgrA deletion mutant z score
2h	0.6547	1.9640	1.9640	1.9640
4h	1.9640	1.9640	1.9640	1.9640
6h	1.9640	1.9640	1.9640	1.9640
72h	1.9640	1.9640	1.9640	1.9640

Acetate

	AgrA z score	AgrB z Score	AgrC z score	AgrA deletion mutant z score
2h	1.0911	1.9640	1.9640	0.6547
4h	1.0911	0.6547	1.9640	1.9640
6h	1.9640	1.0911	1.5275	1.9640
72h	1.9640	0.6547	1.9640	1.9640

Butanol

	AgrA z score	AgrB z Score	AgrC z score	AgrA deletion mutant z score
2h	1.9640	1.9640	1.9640	1.9640
4h	0.6547	1.9640	1.9640	1.9640
6h	1.5275	1.9640	1.9640	1.9640
72h	0.2182	0.2182	1.0911	1.9640

Butyrate

	AgrA z score	AgrB z Score	AgrC z score	AgrA deletion mutant z score
2h	1.0911	1.5275	1.9640	1.9640
4h	1.5275	0.6547	1.9640	1.5275
6h	1.9640	1.0911	1.0911	1.9640
72h	0.2182	1.0911	1.5275	1.9640

Appendix III - Clusters of Orthologous Groups of proteins (COGs)

J - Translation, ribosomal structure and biogenesis

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
661.4	CAC2548	reductase/isomerase/elongation factor domain-containing protein	-	COG5256J
6.1	CAC2976	aspartyl/glutamyl-tRNA amidotransferase subunit B	gatB	COG0064J
3.5	CAC2977	glutamyl-tRNAGln amidotransferase subunit A	gatA	COG0154J
2.8	CAC3056	nucleoside-diphosphate-sugar pyrophosphorylase	-	COG1208 MJ
2.4	CAC3217	peptidyl-tRNA hydrolase	spoVC	COG0193J
Early exponential genes more highly expressed in the wildtype				
3.3	**CAC3423**	**acetyltransferase**	**-**	**COG1670J**
Late exponential genes more highly expressed in the *agrB* mutant				
12.3	CAC1603	hypothetical protein CAC1603	-	COG1514J
5.9	CAC0626	tryptophanyl-tRNA synthetase	trpS	COG0180J
5.4	CAC2979	aspartyl-tRNA synthetase	aspC	COG0017J
3.9	CAC2991	methionyl-tRNA synthetase	metG	COG0073R COG0143J
3.3	CAC1448	tetracycline resistance protein	tetP/tetQ	COG0480J
3.3	CAC1661	secreted nucleic acid binding protein	-	COG0222J
3.1	CAC1041	arginyl-tRNA synthetase	argS	COG0018J
3.0	CAC2740	histidyl-tRNA synthetase	hisS	COG0124J
2.9	CAC3010	ATP-dependent RNA	-	COG0513L KJ
2.8	CAC0630	peptide chain ralease factor 3	-	COG4108J
2.8	CAC2978	Glu-tRNA amidotransferase, subunit C	gatC	COG0721J
2.5	CAC0784	ATP-dependent RNA helicase	-	COG0513L KJ
2.5	CAC1256	ribonuclease G/E family protein	-	COG1530J
2.5	CAC3210	hypothetical protein CAC3210	-	COG1188J
2.4	CAC2362	threonyl-tRNA synthetase	thrS	COG0441J
2.3	CAC2671	aspartyl/glutamyl-tRNA amidotransferase subunit C	gatC	COG0721J
2.2	CAC3738	ribonuclease P	rnpA	COG0594J
Late exponential genes more highly expressed in the wildtype				
6.0	CAC3038	isoleucyl-tRNA synthetase	ileS	COG0060J

6.0	CAC3125	ribosomal protein L29	rpmC	COG0255J
5.9	CAC3123	50S ribosomal protein L14	rplN	COG0093J
5.5	CAC3129	30S ribosomal protein S19	rpsS	COG0185J
5.3	CAC3114	50S ribosomal protein L15	rplO	COG0200J
5.3	CAC3122	50S ribosomal protein L24	rplX	COG0198J
5.1	CAC3121	50S ribosomal protein L5	rplE	COG0094J
5.0	CAC3130	50S ribosomal protein L2	rplB	COG0090J
4.9	CAC3128	50S ribosomal protein L22	rplV	COG0091J
4.9	CAC3124	30S ribosomal protein S17	rpsQ	COG0186J
4.8	CAC3126	50S ribosomal protein L16	rplP	COG0197J
4.6	CAC3118	50S ribosomal protein L6	rplF	COG0097J
4.5	CAC3127	30S ribosomal protein S3	rpsC	COG0092J
4.4	CAC3119	30S ribosomal protein S8	rpsH	COG0096J
4.3	CAC3116	30S ribosomal protein S5	rpsE	COG0098J
4.2	CAC3105	30S ribosomal protein S4	rpsD	COG0522J
4.2	CAC3120	30S ribosomal protein S14	rpsN	COG0199J
3.9	CAC3115	50S ribosomal protein L30	rpmD	COG1841J
3.9	CAC3103	50S ribosomal protein L17	rplQ	COG0203J
3.8	CAC3117	50S ribosomal protein L18	rplR	COG0256J
3.6	CAC3107	30S ribosomal protein S13	rpsM	COG0099J
3.6	CAC3106	30S ribosomal protein S11	rpsK	COG0100J
3.5	CAC2076	rRNA methylase	-	COG1189J
3.1	CAC1297	N-terminal fragment of elongation factor Ts	-	COG0264J
2.7	CAC3109	translation initiation factor IF-1	infA	COG0361J
2.7	CAC2845	peptide chain release factor 2	prfB	COG1186J
2.5	CAC3110	ribosomal protein L14E (/ type)	-	COG2163J
2.5	CAC1286	Fe-S oxidoreductase	-	COG0621J
2.4	CAC3136	elongation factor Tu	tuf	COG0050J
2.4	CAC3138	elongation factor G	fus	COG0480J
2.3	CAC1266	pseudouridylate synthase	-	COG0564J
2.2	CAC3146	50S ribosomal protein L10	rplJ	COG0244J

K - Transcription

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
1334.0	CAC1770	putative RNA polymerase sigma factor SigI	-	COG1191K
965.6	CAC1463	transcriptional regulator	-	COG1476K
578.2	CAC2605	TetR/AcrR family transcriptional regulator	-	COG1309K
544.3	CAC2616	iron-dependent transcription repressor	-	COG1321K
342.5	CAC1071	AcrR family transcriptional regulator	-	COG1309K
124.5	CAC1771	hypothetical protein CAC1771	-	COG0085K

17.3	CAC1946	Phage related transcriptional regulator (Xre family)	-	COG1396K
6.0	CAC1945	Phage related anti-repressor protein	-	COG3561K
3.0	CAC1516	response regulator	-	COG0745TK
3.0	CAC2435	response regulator	-	COG0745TK
3.0	CAC2430	transcription elongation factor, greA	greA	COG0782K
2.6	CAC3046	LytR family transcriptional regulator	-	COG1316K
2.5	CAC2759	response regulator	-	COG0745TK
2.4	CAC2768	AcrR family transcriptional regulator	-	COG1309K
2.4	CAC1451	AraC-type DNA-binding domain-containing protein	-	COG2207K
2.4	CAC1536	AcrR family transcriptional regulator	-	COG1309K
2.4	CAC1032	transcriptional regulator	-	COG1522K
2.3	CAC0865	two-component response regulator	-	COG2197TK
2.2	CAC0564	response regulator	-	COG0745TK
2.1	CAC0422	transcriptional antiterminator licT	licT	COG3711K
2.1	CAC1355	BglG family transcriptional antiterminator	-	COG3711K
2.1	CAC2735	response regulator	-	COG0745TK
Early exponential genes more highly expressed in the wildtype				
72.4	CAC0078	putative accessory gene regulator protein	*agrB*	COG4512OTK
36.8	CAC3409	LysR family transcriptional regulator	alsR	COG0583K
15.2	CAC0550	RNA polymerase sigma factor	-	COG1191K
4.0	CAC3063	LytR family transcriptional regulator	-	COG1316K
3.0	CAC3345	AcrR family transcriptional regulator	-	COG1309K
2.3	CAC2259	MarR family transcriptional regulator	-	COG1321K
Late exponential genes more highly expressed in the *agrB* mutant				
6.8	CAC3283	MarR family transcriptional regulator	-	COG1846K
6.2	CAC2471	TetR/AcrR family transcriptional regulator	-	COG1309K
5.9	CAC3649	Possible stage V sporulation protein T, transcriptional regulator AbrB homolog	spoVT	COG2002K
5.8	CAC3496	transcriptional regulator	-	COG1733K
5.8	CAC1536	AcrR family transcriptional regulator	-	COG1309K
5.3	CAC2934	Predicted transcriptional regulator	-	COG1733K
5.1	CAC3200	pantothenate kinase	-	COG1521K
4.8	CAC0876	MarR family transcriptional regulator	-	COG1846K

4.4	CAC3458	hypothetical protein CAC3458	-	COG0789K
4.3	CAC0841	transcriptional regulator	-	COG1396K
4.2	CAC3687	AcrR family transcriptional regulator	-	COG1309K
4.2	CAC0465	dicA/hipB/ansR family transcriptional regulator	-	COG1396K
4.2	CAC0531	RpiR family transcriptional regulator	-	COG1737K
3.9	CAC3525	hypothetical protein CAC3525	-	COG1959K
3.7	CAC0493	hypothetical protein CAC0493	-	COG0864K
3.5	CAC1950	hypothetical protein CAC1950	-	COG3655K
3.4	CAC0457	AcrR family transcriptional regulator	-	COG1309K
3.4	CAC0023	LysR family transcriptional regulator	-	COG0583K
3.3	CAC0371	response regulator	-	COG0745TK
3.3	CAC3192	transciptional regulator CTSR	-	COG4463K
3.1	CAC1947	phage related transcriptional regulator	-	COG1396K
3.0	CAC3553	LacI family transcription regulator	-	COG1609K
3.0	CAC3338	Predicted transcriptional regulator	-	COG1733K
2.8	CAC1578	transcriptional regulator	-	COG1396K
2.8	CAC2608	AraC family transcriptional regulator	-	COG2207K
2.6	CAC0957	Xre family DNA-binding domain-/TPR repeat-containing protein	-	COG0457R COG1396K
2.6	CAC0212	biotin-(acetyl-CoA carboxylase) ligase	-	COG0340H COG1654K
2.5	CAC2568	transcriptional regulator	-	COG1733K
2.4	CAC3198	transcription elongation factor GreA	greA	COG0782K
2.4	CAC2473	transcriptional regulator	-	COG1476K
2.4	CAC0382	transcriptional regulator	levR	COG1221KT COG3933K
2.4	CAC3438	BlaI/MecI subfamily transcriptional regulator	-	COG3682K
2.3	CAC1786	transcriptional repressor CodY	codY	COG4465K
2.0	CAC1675	transcriptional regulator	-	COG1959K
2.0	CAC1467	Lpr family transcriptional regulator	-	COG1522K
2.0	CAC2768	AcrR family transcriptional regulator	-	COG1309K
Late exponential genes more highly expressed in the wildtype				
26496.0	CAC3409	LysR family transcriptional regulator	alsR	COG0583K
8621.0	CAC1696	sporulation sigma factor SigG	sigG	COG1191K
3802.0	CAC3063	LytR family transcriptional regulator	-	COG1316K
2184.0	CAC1509	specialized sigma subunit of RNA polymerase	-	COG1595K
1849.0	CAC3046	LytR family transcriptional regulator	-	COG1316K
1272.0	CAC0078	putative accessory gene regulator protein	*agrB*	COG4512OTK

3.6	CAC3104	DNA-directed RNA polymerase subunit alpha	rpoA	COG0202K
3.1	CAC2074	arginine repressor	argR	COG1438K
2.1	CAC1668	AcrR family transcriptional regulator	-	COG1309K

D - Cell division and chromosome partitioning

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
654.3	CAC1981	hypothetical protein CAC1981	-	COG0455D
12.2	CAC1161	ATPase of HSP70 class	-	COG0849D
6.3	CAC1871	FtsK-like DNA segregation ATPase	-	COG1674D
6.0	CAC0408	DNA segregation ATP-ase FtsK/SpoIIIE	-	COG1674D
Early exponential genes more highly expressed in the wildtype				
9.4	CAC0071	hypothetical protein CAC0071	-	COG2846D
3.4	CAC3060	CPSC/CAPB subfamily ATPase	-	COG0489D
Late exponential genes more highly expressed in the *agrB* mutant				
3.3	CAC1586	hypothetical protein CAC1586	-	COG0239D
2.6	CAC0497	cell division ATP-binding protein	ftsE	COG2884D
2.5	CAC1249	septum site-determining protein MinD, ATPase	minD	COG2894D
2.4	CAC2145	chromosome partitioning ATPase	-	COG0455D
2.2	CAC1587	hypothetical protein CAC1587	-	COG0239D
2.2	CAC1251	cell cycle protein FtsW	rodA	COG0772D
Late exponential genes more highly expressed in the wildtype				
207.4	CAC1981	hypothetical protein CAC1981	-	COG0455D
204.7	CAC1871	FtsK-like DNA segregation ATPase	-	COG1674D
36.7	CAC3060	CPSC/CAPB subfamily ATPase	-	COG0489D
6.2	CAC0408	DNA segregation ATP-ase FtsK/SpoIIIE	-	COG1674D
5.2	CAC3040	CPSC/CAPB subfamily ATPase	-	COG0489D
3.8	CAC0071	hypothetical protein CAC0071	-	COG2846D
3.4	CAC1242	hypothetical protein CAC1242	mreB	COG1077D
3.0	CAC3733	tRNA uridine 5-carboxymethylaminomethyl modification enzyme GidA	gidA	COG0445D
2.8	CAC2861	sporulation protein SpoIID	spoIID	COG2385D
2.1	CAC2118	cell division protein DivIVA	-	COG3599D
2.0	CAC1248	septum formation inhibitor	minC	COG0850D

L - DNA replication, recombination and repair

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
667.6	CAC1143	exodeoxyribonuclease V, Alpha subunit, RecD	recD	COG0507L
593.9	CAC3514	Reverse transcriptase/maturase family protein	-	COG3344L
4.8	CAC0404	TPR repeat-containing serine/threonin protein kinase	-	COG0515R TKL
3.2	CAC0738	DNA polymerase III subunit epsilon	-	COG0847L
2.7	CAC1763	hypothetical protein CAC1763	-	COG0792L
2.3	CAC1721	primosome assembly protein PriA	rpiA	COG1198L
2.2	CAC2550	hypothetical protein CAC2550	-	COG1533L
2.1	CAC2232	ssDNA exonuclease, RecJ	recJ	COG0608L
Early exponential genes more highly expressed in the wildtype				
4.4	CAC0326	recombination factor protein RarA	-	COG2256L
2.9	CAC2382	single-strand DNA-binding protein	ssb	COG0629L
2.8	CAC0502	excinuclease ABC subunit B	uvrB	COG0556L
2.8	CAC0503	excinuclease ABC subunit A	uvrA	COG0178L
Late exponential genes more highly expressed in the *agrB* mutant				
12.1	CAC1397	hydrolase of PHP superfamily	-	COG0084L
7.4	CAC2276	SS-DNA-specific exonuclease domain-containing protein	-	COG0608L
5.8	CAC3261	methylated DNA-protein cysteine methyltransferase	ada	COG0350L
4.5	CAC1501	DNA-methyltransferase (cytosine-specific), ortholog of BSP6I Bsubtilis	-	COG0270L
4.4	CAC1595	XerC/XerD family integrase/recombinase	-	COG0582L
3.8	CAC1195	NAD-dependent DNA ligase LigA	ligA	COG0272L
3.6	CAC0978	elongation subunit of DNA-dependent DNA polymerase	-	COG3359L
3.2	CAC1951	site-specific recombinase	-	COG1961L
3.0	CAC1785	DNA topoisomerase I	topA	COG0550L COG0551L
2.9	CAC0945	single-stranded DNA-binding protein	ssb	COG0629L
2.5	CAC1273	DNA polymerase III subunit delta	holA	COG1466L
Late exponential genes more highly expressed in the wildtype				
1960.0	CAC0656	transposase (3' fragment)	-	COG3666L
1424.0	CAC2854	ATP-dependent exoDNAse (exonuclease V), alpha subunit, RecD	recD	COG0507L
130.2	CAC0657	transposase (5' fragment)	-	COG3666L
87.5	CAC1628	DNA topoisomerase IV subunit A	gyrA	COG0188L
7.1	CAC2382	single-strand DNA-binding protein	ssb	COG0629L
4.6	CAC1609	Zn-finger containing protein	-	COG1290C COG1933L

O - Posttranslational modification, protein turnover, chaperones

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the wildtype				
136.8	CAC3412	protein-S-isoprenylcysteine methyltransferase	-	COG2020O
35.3	CAC1421	pyruvate-formate lyase-activating enzyme	-	COG1180O
26.3	CAC1765	ATPase, competence protein ComM	-	COG0606O
5.5	CAC2135	ATP-dependent serine protease	-	COG1067O
5.2	CAC0742	phosphatase domain-containing protein	-	COG0671I COG1404O
3.5	CAC2433	HtrA-like serine protease	-	COG0265O
3.5	CAC3218	trypsin-like serine protease	-	COG0265O
3.2	CAC2638	Lon-like ATP-dependent protease	lonB	COG1067O
3.0	CAC0625	periplasmic aspartyl protease	-	COG4934O
3.0	CAC0809	hydrogenase formation factor (hypE)	hypE	COG0309O
2.9	CAC0808	hydrogenase expression factor (hybG)	hybG	COG0298O
2.1	CAC1541	hypothetical protein CAC1541	-	COG0542O
2.1	CAC3549	subtilisin-like serine protease	-	COG1404O
Early exponential genes more highly expressed in the wildtype				
8.8	CAC0557	Zn-dependent protease with chaperone function	-	COG0501O
4.8	CAC1548	thioredoxin reductase	trxB	COG0492O
3.9	CAC1570	glutathione peroxidase	bsaA	COG0386O
3.5	CAC1571	glutathione peroxidase	-	COG0386O
3.2	CAC1549	glutathione peroxidase	bsaA	COG0386O
2.6	CAC3625	MoaA family Fe-S oxidoreductase	-	COG0602O
2.4	CAC3306	Thiol peroxidase, Tpx	-	COG2077O
2.3	CAC2206	flagellar protein FliS	fliS	COG1516N UO
2.0	CAC3082	thioredoxin reductase	-	COG0492O
Late exponential genes more highly expressed in the *agrB* mutant				
5.0	CAC3288	Iron-regulated ABC transporter ATPase	-	COG0396O
3.6	CAC0808	hydrogenase expression factor (hybG)	hybG	COG0298O
3.6	CAC2245	stress-induced protein OsmC	-	COG1765O
3.3	CAC2341	collagenase family protease	-	COG0826O
3.3	CAC3289	Iron-regulated ABC-type transporter membrane component (SufB)	-	COG0719O
3.3	CAC2433	HtrA-like serine protease	-	COG0265O
3.2	CAC3290	Iron-regulated ABC-type transporter membrane component (SufB)	-	COG0719O

3.2	CAC1570	glutathione peroxidase	bsaA	COG0386O
3.1	CAC3367	hypothetical protein CAC3367	-	COG0459O
2.3	CAC0327	hypothetical protein CAC0327	-	COG1225O
2.2	CAC3189	ABC transporter ATPase	clpC	COG0542O
2.1	CAC1571	glutathione peroxidase	-	COG0386O
2.1	CAC2839	hypothetical protein CAC2839	-	COG1214O
2.1	CAC1281	molecular chaperone GrpE	grpE	COG0576O
2.1	CAC0602	ATP-dependent zinc metallopeptidase FtsH (cell dividion protein)	ftsH	COG0465O
2.0	CAC0869	thioredoxine reductase	-	COG0492O
2.0	CAC3714	molecular chaperone (small heat shock protein), HSP18	-	COG0071O
Late exponential genes more highly expressed in the wildtype				
717.3	CAC1765	ATPase, competence protein ComM	-	COG0606O
690.4	CAC0811	hydrogenase expression-formation factor (hypD)	-	COG0409O
564.2	CAC3549	subtilisin-like serine protease	-	COG1404O
511.3	CAC3381	membrane protease subunit stomatin/prohibitin-like protein	-	COG0330O
14.7	CAC0742	phosphatase domain-containing protein	-	COG0671I COG1404O
7.6	CAC1421	pyruvate-formate lyase-activating enzyme	-	COG1180O
4.1	CAC1541	hypothetical protein CAC1541	-	COG0542O
4.0	CAC0463	Serine protease Do (heat-shock protein)	-	COG0265O
3.3	CAC3625	MoaA family Fe-S oxidoreductase	-	COG0602O
2.7	CAC2638	Lon-like ATP-dependent protease	lonB	COG1067O
2.3	CAC3716	Lon-like ATP-dependent protease	lonB	COG1067O
2.1	CAC2777	glutaredoxin	-	COG0695O

M - Cell envelope biogenesis, outer membrane

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
1378.0	CAC1702	hypothetical protein CAC1702	-	COG2148M
89.4	CAC2057	D-alanyl-D-alanine carboxypeptidase	-	COG1686M
83.7	CAC1615	glycosyltransferase	-	COG0463M
Early exponential genes more highly expressed in the wildtype				
6.2	CAC2944	integrin-like repeat-containing cell wall-associated hydrolase	-	COG0791M
4.4	CAC3072	mannose-1-phosphate guanylyltransferase	-	COG0836M
4.3	CAC3068	glycosyltransferase	-	COG0438M
4.3	CAC3062	PHP family hydrolase	-	COG4464GM
4.1	CAC3070	glycosyltransferase	-	COG0438M
4.1	CAC3073	sugar transferase involved in	-	COG2148M

		lipopolysaccharide synthesis		
4.0	CAC3071	glycosyltransferase	-	COG0438M
3.8	CAC3064	UDP-N-acetylglucosamine 2-epimerase	-	COG0381M
3.6	CAC2695	metallo-dependent hydrolase	-	COG3409M
3.4	CAC3066	glycosyltransferase	-	COG0438M
3.4	CAC2523	glycosyltransferase	-	COG1215M
3.2	CAC2663	cell-wall hydrolase domain-containing protein	-	COG0039C COG0791M
2.2	CAC2943	integrin-like repeat-containing cell wall-associated hydrolase	-	COG0791M
2.1	CAC2895	D-alanine-D-alanine ligase	ddlA	COG1181M
Late exponential genes more highly expressed in the *agrB* mutant				
3.0	CAC2776	hypothetical protein CAC2776	-	COG3209M
2.9	CAC3600	dihydrodipicolinate synthase	dapA	COG0329E M
2.6	CAC3250	glutamate racemase	-	COG0796M
2.5	CAC1267	D-alanyl-D-alanine carboxypeptidase	dacB	COG1686M
2.1	CAC2895	D-alanine-D-alanine ligase	ddlA	COG1181M
Late exponential genes more highly expressed in the wildtype				
3106.0	CAC2906	spore coat protein cotS related	-	COG0510M
2498.0	CAC3045	PHP family hydrolase	-	COG4464G M
678.5	CAC3081	spore-cortex-lytic enzyme, SLEB	-	COG3773M
589.6	CAC2186	glycosyltransferase	-	COG3980M
518.8	CAC1564	Closely related to alginate O-acetylation protein (algI)	-	COG1696M
349.2	CAC1876	N-acetylmuramidase	lyc	COG3409M, COG3757M
11.3	CAC2944	integrin-like repeat-containing cell wall-associated hydrolase	-	COG0791M
4.9	CAC1246	penicillin-binding protein 2	pbpA	COG0768M
4.5	CAC2663	cell-wall hydrolase domain-containing protein	-	COG0039C COG0791M
2.5	CAC2063	D-alanyl-D-alanine carboxypeptidase	dacF	COG1686M
2.2	CAC2326	diverged glycosyltransferase domain-containing protein	-	COG0438M
2.0	CAC2127	phospho-N-acetylmuramoyl-pentapeptide-transferase	mraY	COG0472M

N - motility & secretion

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
2.2	CAC1389	ChW repeat-containing and cell-adhesion domain-containing protein	-	COG5492N
Early exponential genes more highly expressed in the wildtype				
15.0	CAC0552	cell adhesion domain-containing protein	-	COG5492N
7.9	CAC0551	cell-adhesion domain-containing protein	-	COG5492N
7.7	CAC2107	cell adhesion domain-containing protein	-	COG5492N
4.6	CAC3086	cell adhesion domain-containing protein	-	COG5492N
4.5	CAC3085	TPR repeat-containing cell adhesion protein	-	COG5492N
3.4	CAC1634	flagellin	-	COG1344N
3.3	CAC3279	ChW repeat-containing protein	-	COG4886S, COG5492N
2.5	CAC3273	ChW repeat-containing protein	-	COG4886S, COG5492N
2.1	CAC3565	cell adhesion domain-containing protein	-	COG5492N
Late exponential genes more highly expressed in the *agrB* mutant				
5.7	CAC0304	chemotaxis motility protein A, gene motA	motA	COG1291N
4.2	CAC2165	flagellar basal body rod protein FlgB	flgB	COG1815N
2.7	CAC2164	flagellar basal body rod protein FlgC	flgC	COG1558N
2.1	CAC2146	flagellar biosynthesis regulator FlhF	flhF	COG1419N
Late exponential genes more highly expressed in the wildtype				
961.8	CAC3274	ChW repeat-containing protein	-	COG4886S, COG5492N
3.5	CAC3565	cell adhesion domain-containing protein	-	COG5492N

P - Inorganic ion transport and metabolism

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
1053.0	CAC0621	exopolyphosphatase	-	COG0248FP
729.9	CAC2800	CotJC-like protein	-	COG3546P
596.2	CAC2540	Co/Zn/Cd efflux system component	-	COG1230P
5.0	CAC3637	oligopeptide ABC transporter, permease component	-	COG1173EP
4.6	CAC3636	oligopeptide ABC transporter, ATPase component	-	COG0444EP
4.3	CAC3643	oligopeptide ABC transporter,	-	COG1173E

		permease component		P
4.2	CAC0682	ammonium transporter	nrgA	COG0004P
2.2	CAC3666	membrane transporter	-	COG2076P
Early exponential genes more highly expressed in the wildtyp3				
7.6	CAC0069	iron-binding protein, hemerythrin	-	COG2703P
2.5	CAC2020	molybdopterin biosynthesis protein MoeA	moeA	COG0303H, COG1910P
2.4	CAC1269	cation transport ATPase	-	COG0474P
Late exponential genes more highly expressed in the *agrB* mutant				
7.6	CAC0069	iron-binding protein, hemerythrin	-	COG2703P
2.5	CAC2020	molybdopterin biosynthesis protein MoeA	moeA	COG0303H, COG1910P
2.4	CAC1269	cation transport ATPase	-	COG0474P
Late exponential genes more highly expressed in the wildtype				
11288.0	CAC1988	ferrichrome-binding periplasmic protein	-	COG0614P
8182.0	CAC1990	ABC-type iron (III) transport system, permease component	-	COG0609P
1445.0	CAC2442	hemin permease	-	COG0609P
950.3	CAC1707	permease component of ATP-dependent phosphate uptake system	-	COG0581P
440.2	CAC1512	formate/nitrate transporter	nirC	COG2116P
4.2	CAC0110	GTPase, sulfate adenylate transferase subunit 1	cysN	COG2895P
4.0	CAC3629	oligopeptide ABC transporter, ATPase component	oppD	COG0444EP
3.7	CAC3630	oligopeptide ABC transporter, permease component	oppC	COG1173EP
2.3	CAC3101	cobalt transporter ATP-binding subunit	-	COG1122P
2.0	CAC0108	sulfate ABC transporter permease	-	COG0600P

T - Signal transduction mechanisms

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
1292.0	CAC2763	methyl-accepting chemotaxis protein	-	COG0840NT
1184.0	CAC3397	methyl-accepting chemotaxis protein	-	COG0840NT
1102.0	CAC1507	sensory transduction histidine kinase	phoR	COG0642T
762.0	CAC2958	PTS system enzyme IIBC component (galactitol/fructose specific)	-	COG1762GT
5.4	CAC0406	FHA domain-containing protein	-	COG1716T
4.7	CAC0317	sensory transduction histidine kinase	-	COG0642T
4.3	CAC0407	PP2C phosphatase family protein	-	COG0631T
4.2	CAC2434	histidine kinase	-	COG0642T
3.6	CAC1553	sensory transduction histidine kinase	-	COG0642T

2.8	CAC0565	histidine kinase	-	COG0642T
2.1	CAC0864	histidine kinase-like ATPase	-	COG4585T
2.1	CAC0111	glutamine-binding periplasmic protein fused to glutamine permease	-	COG0765E, COG0834ET
Early exponential genes more highly expressed in the wildtype				
7.7	CAC0542	methyl-accepting chemotaxis protein	-	COG0840NT
6.6	CAC0815	methyl-accepting chemotaxis protein	-	COG0840NT
5.7	CAC3410	HD-GYP domain (HD superfamily hydrolase)	-	COG2206T
4.2	CAC3352	methyl-accepting chemotaxis protein	-	COG0840NT
4.2	CAC0080	histidine kinase-like ATPase	*agrC*	COG2972T
3.5	CAC0437	sensory transduction histidine kinase	-	COG0642T
3.5	CAC1601	methyl-accepting chemotaxis-like protein (chemotaxis sensory transducer)	-	COG0840NT
3.1	CAC3510	methyl-accepting chemotaxis protein	-	COG0840NT
3.1	CAC3620	amino acid ABC transporter periplasmic-binding protein	-	COG0834ET
3.0	CAC1600	methyl-accepting chemotaxis-like protein (chemotaxis sensory transducer)	-	COG0840NT
2.7	CAC1669	carbon starvation protein	CstA	COG1966T
2.6	CAC1233	chemotaxis protein CheV	chev	COG0784T, COG0835NT
2.5	CAC2217	chemotaxis signal transduction protein CheW	cheW	COG0835NT
2.4	CAC2209	carbon storage regulator, csrA	csrA	COG1551T
2.2	CAC2224	chemotaxis protein CheW	cheW	COG0835NT
2.2	CAC2221	chemotaxis protein methyltransferase, cheR	cheR	COG1352NT
2.1	CAC0122	chemotaxis respons regulator (cheY)	cheY	COG0784T
Late exponential genes more highly expressed in the *agrB* mutant				
5.6	CAC1233	chemotaxis protein CheV	chev	COG0784T, COG0835NT
4.8	CAC0122	chemotaxis respons regulator (cheY)	cheY	COG0784T
3.4	CAC0494	PemK family DNA-binding protein	-	COG2337T
3.4	CAC3437	Zn-dependent protease	-	COG4219KT
3.2	CAC3019	sensory transduction protein	-	COG2199T, COG2200T
3.2	CAC0035	Serine/threonine phosphatase (inactivated protein)	-	COG0631T
3.0	CAC0631	signal transduction protein	-	COG2199T, COG2200T

2.6	CAC2940	membrane-associated sensory histidine kinase-like ATPase	-	COG4585T
2.4	CAC3650	HD-GYP domain-containing protein	-	COG2206T
2.4	CAC1832	LexA repressor	lexA	COG1974KT
2.2	CAC1670	response regulator	-	COG3279KT
2.1	CAC3325	periplasmic amino acid binding protein	-	COG0834ET
Late exponential genes more highly expressed in the wildtype				
4938.0	CAC0118	chemotaxis protein cheA	cheA	COG0643NT
2478.0	CAC3410	HD-GYP domain (HD superfamily hydrolase)	-	COG2206T
2182.0	CAC0119	chemotaxis protein cheW	cheW	COG0835NT
1387.0	CAC0120	membrane-associated methyl-accepting chemotaxis protein	-	COG0840NT
485.1	CAC3662	membrane associated sensory histidine kinase	-	COG0642T
25.3	CAC0903	sensory transduction histidine kinase	-	COG0642T
13.1	CAC2577	GGDEF domain-containing protein	-	COG2199T
2.8	CAC1793	P-loop ATPase domain-containing fused to LexA-like protein	lexA	COG1974KT

C - Energy production and conversion

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
543.4	CAC2498	carbon monoxide dehydrogenase, catalytic subunit (cooS)	-	COG1151C
4.9	CAC3030	hypothetical protein CAC3030	-	COG0716C
3.3	CAC2657	polyferredoxin	-	COG1145C
3.3	CAC0587	flavodoxin	-	COG0716C
2.4	CAC2542	FAD/FMN-containing dehydrogenase	-	COG0277C
2.3	CAC0824	thioredoxin	-	COG0526OC
2.0	CAC1551	nitroreductase family protein	-	COG0778C
Early exponential genes more highly expressed in the wildtype				
25.7	CAC3408	NADH oxidase	-	COG0446R, COG1902C
5.8	CAC2458	2-oxoglutarate ferredoxin oxidoreductase subunit beta	-	COG1013C
5.7	CAC2459	2-oxoacid:ferredoxin oxidoreductase, alpha subunit	-	COG0674C, COG1014C
5.3	CAC2007	glycosyltransferase	-	COG1819GC
5.2	CAC2010	Fe-S oxidoreductase	-	COG1032C
4.4	CAC2018	aldehyde:ferredoxin oxidoreductase	-	COG2414C

4.1	CAC0075	ferredoxin	-	COG1145C
3.7	CAC2026	flavodoxin	-	COG0426C
3.4	CAC2000	indolepyruvate oxidoreductase subunit B	iorB	COG1014C
3.4	CAC2001	indolepyruvate ferredoxin oxidoreductase, subunit alpha	iorA	COG4231C
3.2	CAC1997	glycosyltransferase	-	COG1819GC
2.5	CAC2716	glycosyl transferase	-	COG1819GC
2.3	CAC2575	rubrerythrin	rubY	COG1592C
Late exponential genes more highly expressed in the *agrB* mutant				
14.7	CAC0267	L-lactate dehydrogenase	ldh	COG0039C
7.9	CAC1710	Fe-S oxidoreductase	-	COG1625C
4.3	CAC3527	ferredoxin	-	COG1145C
3.0	CAC0980	pyruvate-formate lyase	pflB	COG1882C
2.9	CAC2830	acylphosphatase AcyP	-	COG1254C
2.8	CAC1656	hypothetical protein CAC1656	-	COG1454C
2.8	CAC0617	hypothetical protein CAC0617	-	COG1413C
2.3	CAC0718	nitroreductase	-	COG0778C
2.0	CAC1621	Fe-S oxidoreductase	-	COG0731C
Late exponential genes more highly expressed in the wildtype				
664.9	CAC0251	hypothetical protein CAC0251	-	COG1600C
210.5	CAC3408	NADH oxidase	-	COG0446R, COG1902C
16.2	CAC2575	rubrerythrin	rubY	COG1592C
3.2	CAC2716	glycosyl transferase	-	COG1819GC
3.1	CAC2844	galactose-1-phosphate uridylyltransferase	galT	COG1085C

G - Carbohydrate transport and metabolism

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
1058.0	CAC0664	sugar-binding periplasmic protein	-	COG1653G
914.8	CAC2956	PTS system enzyme IIC component (galactitol/fructose specific)	-	COG3775G
825.4	CAC0697	galactose mutarotase-like protein	-	COG2017G
708.0	CAC3451	sugar/Na+(H+) simporter	-	COG2211G
648.4	CAC2686	maltodextrin glucosidase	-	COG0366G
605.7	CAC0665	ABC-type sugar transport system, permease component	-	COG1175G
403.3	CAC1342	L-arabinose isomerase	araA	COG2160G
299.7	CAC2179	nucleoside-diphosphate-sugar epimerase	-	COG0451MG
68.0	CAC1359	xylanase/chitin deacetylase	-	COG0726G
4.0	CAC0692	glucuronate isomerase	uxaC	COG1904G
3.7	CAC1351	periplasmic sugar-binding protein	-	COG1879G

2.5	CAC0539	ChW repeat-containing mannanase ManB	manB	COG4124G
2.4	CAC3054	phosphoheptose isomerase	-	COG0279G
2.4	CAC0540	ChW repeat-containing mannanase ManB	-	COG4124G
Early exponential genes more highly expressed in the wildtype				
6.7	CAC2807	endo-1,3(4)-beta-glucanase family protein 16	-	COG2273G
2.4	CAC0712	phosphoglyceromutase	pgm	COG0696G
2.3	CAC2194	nucleoside-diphosphate sugar epimerase	-	COG1086MG
2.2	CAC2166	nucleoside-diphosphate-sugar epimerase	-	COG0451MG
Late exponential genes more highly expressed in the *agrB* mutant				
6.3	CAC1354	phosphotransferase system IIA component	-	COG2190G
3.2	CAC3021	phosphoglycerate mutase	-	COG0406G
2.9	CAC1430	sugar metabolism transcriptional regulator	-	COG1349KG
2.8	CAC1604	methylglyoxal synthase	mgsA	COG1803G
2.8	CAC1404	sugar metabolism transcriptional regulator	-	COG1349KG
2.6	CAC0944	transketolase	tkt	COG0021G
Late exponential genes more highly expressed in the wildtype				
10130.0	CAC0664	sugar-binding periplasmic protein	-	COG1653G
9910.0	CAC2239	glycogen synthase, glgA	glgA	COG0297G
8460.0	CAC0666	sugar permease	-	COG0395G
3444.0	CAC0665	ABC-type sugar transport system, permease component	-	COG1175G
2494.0	CAC0770	glycerol uptake facilitator protein, permease	-	COG0580G
2225.0	CAC1349	aldose-1-epimerase	galM	COG2017G
2001.0	CAC3670	sugar permease	amyC	COG0395G
1241.0	CAC0358	xylanase/chitin deacetylase	-	COG0726G
976.6	CAC3499	endonuclease	-	COG1082G
19.6	CAC2556	endoglucanase family protein	-	COG4124G
6.6	CAC0183	NagC/XylR family transcriptional regulators	-	COG2971G
3.4	CAC2075	kinase	-	COG0061G

E - Amino acid transport and metabolism

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
1233.0	CAC0376	N-dimethylarginine	-	COG1834E
586.5	CAC0852	amino acid permease	-	COG0531E
567.0	CAC2249	asparagine synthase	-	COG0367E

69.0	CAC0737	glutamate dehydrogenase	-	COG0334E
6.6	CAC0940	phosphoribosylformimino-5-aminoimidazole carboxamide ribonucleotide (ProFAR) isomerase	hisA	COG0106E
3.1	CAC3053	histidinol phosphatase related enzyme	-	COG0241E
2.9	CAC0681	nitrogen regulatory protein PII, gene nrgB	nrgB	COG0347E
2.9	CAC0375	aminotransferase A	patA	COG0436E
2.8	CAC1472	amino acid permease	-	COG1113E
2.5	CAC3634	oligopeptide ABC transporter, periplasmic substrate-binding component	-	COG4166E
2.3	CAC0941	imidazoleglycerol-phosphate synthase	hisF	COG0107E
2.2	CAC3285	amino acid transporter	-	COG0531E
Early exponential genes more highly expressed in the wildtype				
3.3	CAC2518	extracellular neutral metalloprotease NPRE	-	COG3227E
Late exponential genes more highly expressed in the *agrB* mutant				
3.8	CAC2338	lysine decarboxylase	-	COG1982E
3.4	CAC2354	Nifs family aminotransferase	-	COG0520E
3.3	CAC2601	S-adenosylmethionine decarboxylase	-	COG1586E
3.0	CAC3448	tesA-like protease	-	COG2755E
2.9	CAC0819	phosphoribosylpyrophosphate synthetase	-	COG0462FE
2.8	CAC2310	Methenyl tetrahydrofolate cyclohydrolase (serine cycle enzyme)	-	COG3404E
2.7	CAC2850	proline/glycine betaine ABC-type transport system, ATPase component	-	COG1125E
2.4	CAC0297	lysine decarboxylase	-	COG1982E
Late exponential genes more highly expressed in the wildtype				
2508.0	CAC3160	indole-3-glycerol-phosphate synthase	trpC	COG0134E
1863.0	CAC0368	4-aminobutyrate aminotransferase	-	COG0160E
978.2	CAC3158	tryptophan synthase subunit beta	trpB	COG0133E
460.5	CAC2719	ethanolamin permease	-	COG0531E
298.9	CAC0737	glutamate dehydrogenase	-	COG0334E
224.3	CAC1369	histidinol-phosphate aminotransferase	hisC	COG0079E
5.0	CAC2517	extracellular neutral metalloprotease, NPRE	nrpE	COG3227E

F - Nucleotide transport and metabolism

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
1140.0	CAC3471	guanosine 5'-monophosphate oxidoreductase	-	COG0516F

2.7	CAC2494	hypothetical protein CAC2494	-	COG0458EF
Early exponential genes more highly expressed in the wildtype				
n/a	n/a	n/a	n/a	n/a
Late exponential genes more highly expressed in the *agrB* mutant				
4.6	CAC0672	Fision threonyl-tRNA synthetase (N-terminal part) and uridine kinase	-	COG0441J, COG0572F
3.5	CAC2113	pyrimidine regulatory protein PyrR Uracil phosphoribosyltransferase	pyrR	COG2065F
2.7	CAC1718	guanylate kinase	gmk	COG0194F
2.6	CAC2654	aspartate carbamoyltransferase catalytic subunit	pyrB	COG0540F
2.5	CAC0650	adenylate cyclase	-	COG1437F
2.5	CAC0872	xanthine permease	-	COG2233F
2.5	CAC2876	deoxycytidylate deaminase	-	COG2131F
2.4	CAC1425	DUTPase, dut	-	COG0756F
2.3	CAC3155	hypothetical protein CAC3155	-	COG1351F
Late exponential genes more highly expressed in the wildtype				
1034.0	CAC3276	ribonucleotide-diphosphate reductase subunit beta	nrdF	COG0208F
4.7	CAC3112	adenylate kinase	adk	COG0563F
2.2	CAC0353	2,3-cyclic-nucleotide 2'phosphodiesterase (duplication)	-	COG0737F
2.1	CAC0027	orotate phosphoribosyltransferase	pyrE	COG0461F
2.1	CAC2064	purine nucleoside phosphorylase	deoD	COG0005F

H - Coenzyme metabolism

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
436.6	CAC1200	putative phospho-adenylylsulfate sulfotransferase	-	COG0175EH
6.9	CAC0591	riboflavin synthase subunit alpha	ribB	COG0307H
5.8	CAC0592	riboflavin biosynthes protein RIBA (gtpcyclohydrolase/3,4-dihydroxy-2-butanone 4-phosphate synthase)	ribA	COG0108H, COG0807H
5.0	CAC0590	riboflavin biosynthesis protein RIBD (pirimidine deaminase and pirimidine reductase)	ribD	COG0117H, COG1985H
4.3	CAC0593	riboflavin synthase subunit beta	ribH	COG0054H
2.4	CAC0091	ketol-acid reductoisomerase	ilvC	COG0059EH
Early exponential genes more highly expressed in the wildtype				
4.1	CAC2006	enzyme of siderophore/surfactin biosynthesis	-	COG0736I, COG2091H
3.7	CAC2021	molybdopterin biosynthesis protein MoeA	moeA	COG0303H
2.6	CAC1994	molybdopterin biosynthesis protein MoaB	moaB	COG0521H
2.6	CAC2022	molybdopterin biosynthesis protein MoaB	moaB	COG0521H

2.2	CAC0584	precorrin-6B methylase 1 CobL1/CbiE	-	COG2241H
Late exponential genes more highly expressed in the *agrB* mutant				
n/a	n/a	n/a	n/a	n/a
Late exponential genes more highly expressed in the wildtype				
6827.0	CAC1993	molybdenum cofactor biosynthesis protein MoaA	moaA	COG2896H
3658.0	CAC1376	precorrin-8X methylmutase	cbiC	COG2082H
2140.0	CAC1381	cobalt-precorrin-6x reductase	cbiJ/cobK	COG2099H
1528.0	CAC1375	cobyrinic acid a,c-diamide synthase CobB	cobB	COG1797H
1137.0	CAC1382	precorrin-3B C17-methyltransferase	cbiH/cobJ	COG1010H
493.2	CAC1994	molybdopterin biosynthesis protein MoaB	moaB	COG0521H
102.0	CAC1377	cobalt-precorrin-6A synthase	cbiD	COG1903H
36.5	CAC1374	cobyric acid synthase	cbiP	COG1492H
2.7	CAC0109	sulfate adenylyltransferase subunit 2	cysD	COG0175EH
2.4	CAC1099	dephospho-CoA kinase	coaE	COG0237H

I - Lipid metabolism

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
616.6	CAC1364	phospholipase D family endonuclease	-	COG1502I
2.3	CAC1489	phosphatidic acid phosphatase	-	COG0671I
2.2	CAC2902	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	ipk	COG1947I
Early exponential genes more highly expressed in the wildtype				
8.5	CAC0814	3-oxoacyl-	-	COG0332I
6.8	CAC2012	enoyl-CoA hydratase	fadB	COG1024I
6.4	CAC2011	3-oxoacyl-ACP synthase III	fabH	COG0332I
5.6	CAC2016	enoyl-CoA hydratase	fadB	COG1024I
5.2	CAC2009	3-hydroxyacyl-CoA dehydrogenase	mmgB	COG1250I
Late exponential genes more highly expressed in the *agrB* mutant				
40.8	CAC0796	CDP-diglyceride synthetase	-	COG0575I
3.7	CAC2246	lysophospholipase L2 PLDB	-	COG2267I
3.4	CAC1658	phospholipase D family protein	-	COG1502I
2.2	CAC3591	Acyl-ACP thioesterase	-	COG3884I
2.0	CAC0676	phosphatidylserine synthase	pssA	COG1183I
Late exponential genes more highly expressed in the wildtype				
741.0	CAC0799	phosphatidylserine decarboxylase	psd	COG0688I
2.5	CAC3568	Acetyl-CoA carboxylase alpha subunit	accA	COG0825I
2.0	CAC3569	Acetyl-CoA carboxylase beta subunit	accD	COG0777I

Q - Secondary metabolites biosynthesis, transport and catabolism

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
n/a	n/a	n/a	n/a	n/a
Early exponential genes more highly expressed in the wildtype				
8.3	CAC2017	Acyl carrier protein	-	COG0236IQ
6.3	CAC2008	3-oxoacyl-(acyl-carrier-protein) synthase	pksF	COG0304IQ
Late exponential genes more highly expressed in the *agrB* mutant				
n/a	n/a	n/a	n/a	n/a
Late exponential genes more highly expressed in the wildtype				
		n/a		

R - General function prediction only

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
1141.0	CAC1423	dehydrogenase	-	COG1028IQR
1051.0	CAC1345	D-xylose-proton symporter	xylT	COG0477GEPR
680.6	CAC2743	permease	-	COG2270R
670.9	CAC1529	glycosyl hydrolase	-	COG3940R
603.6	CAC2822	TPR repeat-containing protein	-	COG0457R
558.7	CAC1530	galactoside permease	lacY	COG0477GEPR
77.9	CAC3230	hydrogenase subunit (ferredoxin)	-	COG4624R
4.2	CAC2658	glutamine synthetase type III	glnA	COG3968R
4.1	CAC1776	HAD family phosphatase	-	COG1011R
3.5	CAC3055	sugar kinase	-	COG2605R
2.8	CAC3674	CBS domain-containing protein	-	COG0517R
2.6	CAC2992	short-chain dehydrogenase	-	COG0300R
2.4	CAC1677	permease	-	COG0628R
2.4	CAC1504	MDR-type permease	-	COG0477GEPR
2.3	CAC1497	hypothetical protein CAC1497	-	COG3541R
2.3	CAC1496	hypothetical protein CAC1496	-	COG3541R
2.3	CAC1234	hypothetical protein CAC1234	pheB	COG4492R
2.2	CAC3047	hypothetical protein CAC3047	-	COG0728R
2.1	CAC1543	lactate dehydrogenase	-	COG1052CHR
2.1	CAC1433	hypothetical protein CAC1433	-	COG0561R
2.0	CAC2502	permease	-	COG0697GER
Early exponential genes more highly expressed in the wildtype				
7.7	CAC2806	Icc family phosphohydrolase		COG1409R

6.4	CAC3343	biotin synthase related domain-containing protein		COG4277R
6.4	CAC2014	esterase		COG0824R
5.7	CAC0816	lipase-esterase related protein		COG0596R
5.3	CAC0440	hypothetical protein CAC0440		COG2823R
5.1	CAC1010	Icc family phosphohydrolase		COG1409R
4.0	CAC3069	glycosyltransferase		COG1216R
3.9	CAC2002	iron-sulfur flavoprotein		COG0655R
3.2	CAC0077	CAAX-like membrane endopeptidase		COG1266R
2.9	CAC3065	O-antigen/teichoic acid transporter		COG2244R
2.8	CAC0458	permease		COG0477GEPR
2.7	CAC0140	permease		COG1277R
2.2	CAC3659	S-adenosylmethionine-dependent methyltransferase		COG0500QR
2.1	CAC0702	Med/BMP family lipoprotein		COG1744R
Late exponential genes more highly expressed in the *agrB* mutant				
8.8	CAC2593	hypothetical protein CAC2593	-	COG1279R
8.8	CAC2921	thiamine biosynthesis protein ThiH	thiH	COG1060HR
8.2	CAC1083	membrane-associated metal-binding protein	-	COG1988R
6.1	CAC2948	ABC transporter ATPase	-	COG0488R
6.1	CAC0460	CBS-domain containing protein	-	COG1253R
6.1	CAC2840	acetyltransferase	-	COG0456R
5.4	CAC0412	TPR repeat-containing protein	-	COG0457R
5.4	CAC3697	hypothetical protein CAC3697	-	COG0477GEPR
4.4	CAC3526	FMN-binding protein	-	COG3576R
4.2	CAC1764	glutamine amidotransferase	-	COG2071R
4.1	CAC1984	permease	-	COG0697GER
4.1	CAC2501	membrane protein, transporter of cations and cationic drugs	-	COG0697GER
4.1	CAC1831	Zn-finger domain-containing protein	-	COG5273R
3.7	CAC2411	zinc-dependent protease	-	COG0612R
3.6	CAC1322	glycerol-3-phosphate dehydrogenase, GLPA	glpA	COG0579R
3.5	CAC1403	zinc-dependent hydrolase	-	COG0491R
3.4	CAC2772	permease	-	COG2252R
3.4	CAC0307	methyltransferase	-	COG0313R
3.3	CAC1433	hypothetical protein CAC1433	-	COG0561R
3.2	CAC3314	nitroreductase family protein	-	COG3560R
3.2	CAC0364	metallobeta-lactamase superfamily hydrolase	-	COG1234R
3.1	CAC2487	acetyltransferase	-	COG3981R
3.1	CAC0522	HAD superfamily hydrolase	-	COG0561R
3.0	CAC0418	HAD family phosphatase	-	COG0546R
2.7	CAC0084	MutT/NUDIX family hydrolase	-	COG0494L

				R
2.7	CAC2423	HAD family phosphatase	-	COG0561R
2.5	CAC0012	dehydrogenase	-	COG0579R
2.5	CAC0905	NAD(FAD)-dependent dehydrogenase	-	COG0446R
2.5	CAC2298	nucleotidyltransferase	-	COG1708R
2.4	CAC0151	SAM dependent methyltransferase	-	COG0500QR
2.4	CAC2607	gluconate 5-dehydrogenase	-	COG1028IQR
2.4	CAC2472	alpha/beta fold family hydrolase	-	COG0596R
2.3	CAC1468	acetyltransferase	-	COG0454KR
2.2	CAC0271	esterase	-	COG0824R
2.2	CAC1858	MoxR-like ATPase	-	COG0714R
2.2	CAC2491	acetyltransferase	-	COG0454KR
2.2	CAC2775	Phosphohydrolase from calcineurin family	-	COG1408R
2.2	CAC1470	2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase	-	COG0596R
2.1	CAC0991	putative intracellular protease	-	COG0693R
2.1	CAC2314	RfbX family O-antigen/teichoic acid exporter membrane protein	-	COG2244R
Late exponential genes more highly expressed in the wildtype				
4323.0	CAC0776	NCAIR mutase (PurE)-related protein	-	COG1691R
3622.0	CAC0184	acetyltransferase	-	COG0454KR
766.3	CAC3431	membrane export protein, related to SecD/SecF protein exporters	-	COG2409R
716.2	CAC1452	MDR-type permease	-	COG0477GEPR
687.5	CAC2614	Beta-phosphoglucomutase	-	COG0637R
516.2	CAC3230	hydrogenase subunit (ferredoxin)	-	COG4624R
6.0	CAC2571	acetyltransferase	-	COG0454KR
5.5	CAC3735	RNA-binding protein Jag, SpoIIIJ-associated	jag	COG1847R
3.7	CAC3402	dipeptidyl aminopeptidase/acylaminoacyl-peptidase related protein	-	COG0596R
3.3	CAC1804	exopolyphosphatase family protein	-	COG0618R
3.0	CAC1829	hypothetical protein CAC1829	-	COG0312R
2.9	CAC0146	related to ABC transporter permease component	-	COG1277R
2.8	CAC0175	N-acetylmuramic acid-6-phosphate etherase	murQ	COG2103R
2.7	CAC2572	aminoglycoside phosphotransferase	-	COG3173R
2.6	CAC1674	small subunit of NADPH-dependent glutamate synthase	gltB	COG0493ER
2.3	CAC1386	zinc-dependent hydrolase	-	COG0491R

S - Function unknown

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
790.7	CAC1324	hypothetical protein CAC1324	-	COG3862S
788.6	CAC2426	hypothetical protein CAC2426	-	COG3580S
5.2	CAC2841	hypothetical protein CAC2841	-	COG3601S
4.1	CAC0410	hypothetical protein CAC0410	-	COG4842S
3.8	CAC0740	hypothetical protein CAC0740	-	COG1550S
3.7	CAC2484	hypothetical protein CAC2484	-	COG4129S
3.0	CAC2436	hypothetical protein CAC2436	-	COG5438S
2.6	CAC2758	hypothetical protein CAC2758	-	COG4086S
2.4	CAC3050	hypothetical protein CAC3050	-	COG2327S
2.1	CAC2386	beta-propeller fold protein	-	COG3391S
2.1	CAC3340	hypothetical protein CAC3340	-	COG2357S
2.0	CAC1537	hypothetical protein CAC1537	-	COG2013S
2.0	CAC0851	hypothetical protein CAC0851	-	COG3384S
Early exponential genes more highly expressed in the wildtype				
22.2	CAC0063	hypothetical protein CAC0063		COG4926S
18.6	CAC0062	hypothetical protein CAC0062		COG4722S
9.3	CAC0555	hypothetical protein CAC0555		COG2717S
7.2	CAC1080	hypothetical protein CAC1080		COG4676S
5.6	CAC0072	hypothetical protein CAC0072		COG4309S
5.0	CAC0055	hypothetical protein CAC0055		COG4332S
4.9	CAC3033	hypothetical protein CAC3033		COG3584S, COG3883S
4.0	CAC0325	hypothetical protein CAC0325		COG1971S
3.9	CAC0073	hypothetical protein CAC0073		COG1357S
3.7	CAC0029	hypothetical protein CAC0029		COG3863S
3.3	CAC0556	hypothetical protein CAC0556		COG2013S
3.3	CAC3658	hypothetical protein CAC3658		COG1285S
2.6	CAC0580	hypothetical protein CAC0580		COG2718S
2.2	CAC2210	flagellar assembly protein FliW		COG1699S
2.2	CAC2193	hypothetical protein CAC2193		COG4641S
2.1	CAC2196	hypothetical protein CAC2196		COG2604S
Late exponential genes more highly expressed in the *agrB* mutant				
19.3	CAC1676	hypothetical protein CAC1676	-	COG3881S
18.0	CAC2841	hypothetical protein CAC2841	-	COG3601S
16.8	CAC1602	diverged CheY-domain-containing protein	-	COG4378S
7.4	CAC0041	hypothetical protein CAC0041	-	COG4842S
6.6	CAC2928	hypothetical protein CAC2928	-	COG3859S
6.1	CAC0040	hypothetical protein CAC0040	-	COG4842S
5.8	CAC1537	hypothetical protein CAC1537	-	COG2013S
5.7	CAC3383	hypothetical protein CAC3383	-	COG1690S

5.0	CAC3284	hypothetical protein CAC3284	-	COG1307S
4.9	CAC1398	hypothetical protein CAC1398	-	COG0011S
4.7	CAC3156	hypothetical protein CAC3156	-	COG1939S
4.3	CAC1491	ATPase	-	COG4185S
3.8	CAC1095	hypothetical protein CAC1095	-	COG2733S
3.4	CAC1657	hypothetical protein CAC1657	-	COG1357S
3.4	CAC2631	hypothetical protein CAC2631	-	COG1376S
3.3	CAC0413	hypothetical protein CAC0413	-	COG4842S
3.2	CAC2513	hypothetical protein CAC2513	-	COG0464O, COG5373S
3.2	CAC3001	hypothetical protein CAC3001	-	COG3877S
3.1	CAC1699	hypothetical protein CAC1699	-	COG1496S
2.8	CAC1388	mannose-6-phosphate isomerase	manC	COG3837S
2.6	CAC1716	hypothetical protein CAC1716	-	COG1561S
2.6	CAC3340	hypothetical protein CAC3340	-	COG2357S
2.5	CAC1635	hypothetical protein CAC1635	-	COG2996S
2.5	CAC2675	hypothetical protein CAC2675	-	COG4496S
2.4	CAC0634	hypothetical protein CAC0634	-	COG4720S
2.2	CAC1049	hypothetical protein CAC1049	-	COG1728S
2.1	CAC1798	hypothetical protein CAC1798	-	COG0779S
2.0	CAC2401	2-hydroxyglutaryl-CoA dehydratase activator	-	COG1924I, COG3580S, COG3581S
Late exponential genes more highly expressed in the wildtype				
2404.0	CAC3676	hypothetical protein CAC3676		COG3786S
2229.0	CAC0352	hypothetical protein CAC0352		COG1434S
1911.0	CAC0062	hypothetical protein CAC0062		COG4722S
1222.0	CAC2557	hypothetical protein CAC2557		COG4994S
317.1	CAC0733	hypothetical protein CAC0733		COG4878S
107.4	CAC3372	hypothetical protein CAC3372		COG2323S
4.5	CAC0746	secreted protease metal-dependent protease		COG4412S
4.4	CAC1606	hypothetical protein CAC1606		COG2454S
3.9	CAC2569	NimC/NimA family protein		COG5015S

V - defense mechanisms

Fold Change	ORF	Function	Gene	COG
Early exponential genes more highly expressed in the *agrB* mutant				
1240.0	CAC0834	integral membrane protein	-	COG0842V
78.4	CAC3609	ABC-type MDR transport system, permease component	-	COG0842V
9.9	CAC3584	permease	-	COG0577V
9.6	CAC3583	permease	-	COG0577V
8.9	CAC3585	ABC-type transporter, ATPase component	-	COG1136V
4.2	CAC0242	permease	-	COG0842V

4.1	CAC0288	ABC transporter, ATP-binding protein	-	COG1131V
4.1	CAC3404	ABC transporter permease	-	COG0842V
2.7	CAC2932	ABC-type MDR transporter, ATPase component	-	COG1131V
2.6	CAC2930	hypothetical protein CAC2930	-	COG0842V
2.4	CAC3405	ABC-type MDR transporter, ATPase component	-	COG1131V
2.3	CAC3403	hypothetical protein CAC3403	-	COG0842V
2.3	CAC1534	permease	-	COG0577V
2.1	CAC0319	ABC transporter ATP-binding protein	-	COG1136V
Early exponential genes more highly expressed in the wildtype				
n/a	n/a	n/a	n/a	n/a
Late exponential genes more highly expressed in the *agrB* mutant				
6.1	CAC3282	ABC-type multidrug/protein/lipid transport system, ATPase component	-	COG1132V
5.1	CAC1444	hypothetical protein CAC1444	-	COG4767V
3.1	CAC3281	ABC-type multidrug/protein/lipid transport system, ATPase component	-	COG1132V
3.0	CAC2393	ABC transporter ATPase	-	COG1132V
Late exponential genes more highly expressed in the wildtype				
1802.0	CAC3615	hypothetical protein CAC3615	-	COG0577V
679.4	CAC0170	permease	-	COG0842V
485.5	CAC3614	hypothetical protein CAC3614	-	COG0577V
21.7	CAC0082	hypothetical protein CAC0082	-	COG4403V
5.8	CAC2808	beta-lactamase class C domain-containing protein	-	COG1680V

Appendix IV - Publications

Journal of Microbiological Methods 80 (2010) 49–55

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

The ClosTron: Mutagenesis in *Clostridium* refined and streamlined

John T. Heap, Sarah A. Kuehne, Muhammad Ehsaan, Stephen T. Cartman, Clare M. Cooksley, Jamie C. Scott, Nigel P. Minton *

BBSRC Sustainable Bioenergy Centre, School of Molecular Medical Sciences, Centre for Biomolecular Sciences, The University of Nottingham, University Park, Nottingham, NG7 2RD, UK

ARTICLE INFO

Article history:
Received 18 October 2009
Accepted 26 October 2009
Available online 3 November 2009

Keywords:
ClosTron
Intron
Clostridium
Mutagenesis
Mutant

ABSTRACT

The recent development of the ClosTron Group II intron directed mutagenesis tool for *Clostridium* has advanced genetics in this genus, and here we present several significant improvements. We have shown how marker re-cycling can be used to construct strains with multiple mutations, demonstrated using FLP/FRT in *Clostridium acetobutylicum*; tested the capacity of the system for the delivery of transgenes to the chromosome of *Clostridium sporogenes*, which proved feasible for 1.0 kbp transgenes in addition to a marker; and extended the host range of the system, constructing mutants in *Clostridium beijerinckii* and, for the first time, in a B1/NAP1/027 'epidemic' strain of *Clostridium difficile*. Automated intron design bioinformatics are now available free-of-charge at our website http://clostron.com; the out-sourced construction of re-targeted intron plasmids has become cost-effective as well as rapid; and the combination of constitutive intron expression with direct selection for intron insertions has made mutant isolation trivial. These developments mean mutants can now be constructed with very little time and effort for the researcher. Those who prefer to construct plasmids in-house are no longer reliant on a commercial kit, as a mixture of two new plasmids provides unlimited template for intron re-targeting by Splicing by Overlap Extension (SOE) PCR. The new ClosTron plasmids also offer blue–white screening and other options for identification of recombinant plasmids. The improved ClosTron system supersedes the prototype plasmid pMTL007 and the original method, and exploits the potential of Group II introns more fully.

1. Introduction

The eubacterial genus *Clostridium* encompasses organisms of both medical and industrial importance. These include the pathogens *Clostridium tetani*, *Clostridium perfringens*, *Clostridium botulinum* (Hatheway, 1990) and *Clostridium difficile* (Brazier, 2008); and the solvent-producers *Clostridium acetobutylicum* (Papoutsakis, 2008) and *Clostridium thermocellum* (Lynd et al., 2005). Research both to counter pathogens and exploit useful strains has been hindered by a history of limited genetic methods, especially for mutagenesis. A modest number of directed mutants of *Clostridium* strains had been reported prior to 2007, almost all constructed by homologous recombination (summarized in Heap et al., 2009a). Of these, most were either 'single-crossover' strains in which an integrated plasmid serves as an insertional mutagen, or 'double crossover' strains in which an introduced alternative allele is exchanged with the wild-type allele. The former type of mutant is inherently unstable, while the latter type has proven difficult to isolate, partly because the use of negative selection markers, commonly relied upon to facilitate allele exchange in other organisms, has not been established in *Clostridium*.

To address the need for reliable mutagenesis, we recently developed the ClosTron (Heap et al., 2007), a Group II intron directed mutagenesis system for *Clostridium*. Bacterial Group II introns are a relatively newly-characterized type of mobile element (Lambowitz and Zimmerly, 2004) which can be used for the directed construction of stable mutants thanks to two key properties. Firstly, intron target specificity is determined mainly by base-pairing between the target site DNA and intron RNA, which can be rationally modified. Secondly, intron mobility requires the presence of an intron-encoded protein (IEP), which can be provided transiently during mutagenesis and subsequently removed to ensure the stability of the strain; a strategy analagous to the stabilization of mini-transposon insertions by removal of the transposase. Like most other Group II intron mutagenesis systems, the ClosTron plasmid pMTL007 contains a mini-intron derivative of the LLtrB intron from *Lactococcus lactis*, ideal for rational intron re-targeting and loss of the IEP gene *ltrA* (Karberg et al., 2001). Other elements of pMTL007 facilitate conjugal transfer of the plasmid into *Clostridium* spp., its subsequent replication and maintenance, expression of the intron and IEP, and specific selection of clones containing an insertion.

The ClosTron system has transformed our research, and we have distributed pMTL007 to many other laboratories, some of which have already begun to publish studies using ClosTron mutants (Emerson et al., 2009; Kirby et al., 2009; Twine et al., 2009; Underwood et al., 2009). Despite its usefulness, the system and procedures we described

* Corresponding author. Tel.: +44 115 84 67458; fax: +44 115 82 32120.
E-mail addresses: john.heap@nottingham.ac.uk (J.T. Heap), nigel.minton@nottingham.ac.uk (N.P. Minton).

0167-7012/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.mimet.2009.10.018

AEM

An *agr* Quorum Sensing System That Regulates Granulose Formation and Sporulation in *Clostridium acetobutylicum*

Elisabeth Steiner, Jamie Scott, Nigel P. Minton, and Klaus Winzer

The Clostridia Research Group, School of Molecular Medical Sciences, Centre for Biomolecular Sciences, University of Nottingham, Nottingham, United Kingdom

The Gram-positive, anaerobic, endospore-forming bacterium *Clostridium acetobutylicum* has considerable biotechnological potential due to its ability to produce solvents as fermentation products, in particular the biofuel butanol. Its genome contains a putative *agr* locus, *agrBDCA*, known in staphylococci to constitute a cyclic peptide-based quorum sensing system. In staphylococci, *agrBD* is required for the generation of a peptide signal that, upon extracellular accumulation, is sensed by an *agrCA*-encoded two-component system. Using ClosTron technology, *agrB*, *agrC*, and *agrA* mutants of *C. acetobutylicum* ATCC 824 were generated and phenotypically characterized. Mutants and wild type displayed similar growth kinetics and no apparent differences in solvent formation under the conditions tested. However, the number of heat-resistant endospores formed by the mutants in liquid culture was reduced by about one order of magnitude. On agar-solidified medium, spore formation was more strongly affected, particularly in *agrA* and *agrC* mutants. Similarly, accumulation of the starch-like storage compound granulose was almost undetectable in colonies of *agrB*, *agrA*, and *agrC* mutants. Importantly, these defects could be genetically complemented, demonstrating that they were directly linked to *agr* inactivation. A diffusible factor produced by *agrBD*-expressing strains was found to restore granulose and spore formation in the *agrB* mutant. Furthermore, a synthetic cyclic peptide, designed on the basis of the *C. acetobutylicum* AgrD sequence, was also capable of complementing the defects of the *agrB* mutant when added exogenously to the culture. Together, these findings support the hypothesis that *agr*-dependent quorum sensing is involved in the regulation of sporulation and granulose formation in *C. acetobutylicum*.

The Gram-positive, endospore-forming bacterium *Clostridium acetobutylicum* is strictly fermentative and gains energy by converting sugars and starch to organic acids and solvents (5, 17). In the past, it has been exploited for the large-scale production of acetone and butanol, but the classical industrial fermentation process is currently considered uneconomical in most countries, despite increasing oil prices (16). Essential for the engineering of more efficient strains and processes is a more thorough understanding of the organism's physiology and metabolism (22).

The fermentation metabolism of *C. acetobutylicum* is complex. During exponential growth in batch culture, acetic and butyric acid are produced, which allows the cells to gain more than 3 ATPs per molecule of glucose. However, continued acid production poses a problem as the pH of the medium decreases and undissociated acids diffuse back into the cells, thereby affecting intracellular pH and leading to proton gradient dissipation. To avoid acid death, cells shift their metabolism to solvent formation: some of the produced acids (mainly butyric acid) are taken up again and, together with remaining sugars, converted to butanol, acetone, and also some ethanol (5, 17). This switch to solvent production coincides with the initiation of the complex developmental program of sporulation. As part of these processes, a starch-like storage compound termed granulose is transitorily formed and accumulates in the cytoplasm (24, 35). Both solventogenesis and sporulation are linked via the transcription factor Spo0A, the master regulator of sporulation: Spo0A is essential for the initiation of sporulation in solventogenic clostridia (as in *Bacillus subtilis*) but also required for high solvent production (8, 34). In contrast to *B. subtilis*, Spo0A in *C. acetobutylicum* is not activated via a complex phosphorelay system. Instead, its phosphorylation state is controlled by multiple orphan histidine kinases through direct interaction (37). The whole cascade of sigma factors that regulates the sporulation process downstream of phosphorylated Spo0A in *B. subtilis* is present in *C. acetobutylicum* (2, 18) although some differences seem to exist (for instance in the activation of σ^F by SpoIIE). However, while general conditions for solventogenesis and sporulation are known, and progress is being made in unraveling the underlying regulatory networks, the chemical or physical nature of the initiating signals still remains to be elucidated.

There is, however, some evidence to suggest that cell-cell communication (quorum sensing) systems are widely distributed in *Closeridium* spp. (43) and that they may play a role in regulating both solventogenesis and sporulation (3, 20, 21). Quorum sensing is a communication mechanism that relies on small, diffusible signal molecules and allows bacteria to coordinate changes in gene expression with cell population density (42). Most Gram-positive quorum sensing systems are based on secreted peptides which can be linear or cyclic and sometimes contain extensive posttranslational modifications (23, 39). Relatively little is known about the operation of such systems in clostridial species, but for the solventogenic *Clostridium sacharoperbutylacetonicum*, an as-yet-unidentified, self-generated signal in the culture supernatant has been reported to induce solvent formation in a "low-solvent" mutant, suggesting that the metabolic shift from acid to solvent production might be subject to quorum sensing control (20). Furthermore, *in silico* analysis suggests that *agr*-type quorum sensing

Received 29 July 2011 **Accepted** 6 December 2011
Published ahead of print 16 December 2011
Address correspondence to Klaus Winzer, klaus.winzer@nottingham.ac.uk, or Nigel P. Minton, nigel.minton@nottingham.ac.uk.
Supplemental material for this article may be found at http://aem.asm.org/.

doi:10.1128/AEM.06376-11